

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:29:36 ; Search time 142.364 Seconds  
(without alignments)  
1413.528 Million cell updates/sec

Title: US-10-618-320A-1  
Perfect score: 2400  
Sequence: 1 MGICYSLRPLLFGPGDDPC.....VFNDCRDIQRMHLKQYELL 458

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2400	100.0	458	8	ADG74722 Human G-p
2	2400	100.0	458	9	AEAL7292 Human XLG
3	2124	88.5	448	8	ADG74746 Mouse G-p
4	2113	88.0	450	8	ADG74747 Rat G-pro
5	1819	75.8	381	5	ABBO9272 G protein
6	1819	75.8	381	7	ADCO9607 Human G-p
7	1819	75.8	381	7	ADEG1907 Human Pro
8	1819	75.8	381	8	ADU60726 Human G-p
9	1819	75.8	381	9	ADX26261 Novel cel
10	1819	75.8	381	9	AEAL7294 Human Gol
11	1811	75.5	381	7	ADEG1905 Rat Prote
12	1559	65.0	756	5	ABG60299 Lymphona
13	1559	65.0	756	6	ABP97657 Amino aci
14	1559	65.0	909	8	ADQ26060 Guanine n
15	1559	65.0	909	8	ABM82265 Tumour-as
16	1559	65.0	909	9	ADX06936 Cycilin-de
17	1540	64.2	379	4	AAB99060 Human G-p
18	1540	64.2	379	5	ABBO9269 G protein
19	1540	64.2	379	7	ADCO9604 Human G-p
20	1540	64.2	379	7	ADJ68299 Human hea
21	1540	64.2	379	8	ADU60723 Human G-p
22	1537	64.0	720	6	ABP56694 GKRI:Gs f
23	1536.5	64.0	755	8	ADM79379 Mouse lym
24	1529.5	63.7	380	3	AAB23382 Human G-a

25	1529.5	63.7	380	4	AAB99058 Human G-p
26	1529.5	63.7	380	4	AAB99061 Human G-p
27	1529.5	63.7	380	5	ABBO9270 G protein
28	1529.5	63.7	380	7	ADCO9605 Human G-p
29	1529.5	63.7	380	7	ADP70778 Minicell
30	1529.5	63.7	380	8	ADQ26061 Guanine n
31	1529.5	63.7	380	8	ABM82267 Tumour-as
32	1529.5	63.7	380	8	ADU60724 Human G-p
33	1526.5	63.6	926	4	AAU04387 GPCR-Gs f
34	1526.5	63.6	926	7	ADL96550 G protein
35	1526.5	63.6	926	9	ADM44723 Human RUP
36	1526.5	63.6	926	9	AEB20907 Human RUP
37	1525.5	63.6	394	2	AAR94559 Human Gs
38	1525.5	63.6	394	5	ABBO9267 G protein
39	1525.5	63.6	394	5	ABG60304 Lymphona
40	1525.5	63.6	394	6	ABP97662 Amino aci
41	1525.5	63.6	394	7	ABR82636 C. elegan
42	1525.5	63.6	394	7	ADC09602 Human G-p
43	1525.5	63.6	394	7	ADP70779 Minicell
44	1525.5	63.6	394	8	ADQ26059 Guanine n
45	1525.5	63.6	394	8	ABM82266 Tumour-as

## ALIGNMENTS

RESULT 1  
ADG74722  
ID ADG74722 standard; protein; 458 AA.  
XX  
AC ADG74722;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Human G-protein Gml amino acid sequence.  
XX  
KW G protein; Gml; G protein-coupled receptor mediated signal transduction;  
KW GTP binding site; GTPase site; G protein alpha subunit;  
KW signal transduction; G-protein-coupled receptor.  
XX  
OS Homo sapiens.  
XX  
PN EP1382613-A1.  
XX  
PD 21-JAN-2004.  
XX  
PF 09-JUL-2003; 2003EP-00015519.  
XX  
PR 16-JUL-2002; 2002JP-00206841.  
PR 19-DEC-2002; 2002JP-00367778.  
PR 31-MAR-2003; 2003JP-00095955.  
XX  
PA (SUMO ) SUMITOMO CHEM CO LTD.  
XX  
PI Takahashi Y, Matsumoto Y, Oeda K;  
DR WPI: 2004-111483/12.  
DR N-PSDB; ADG74723.  
XX  
PT New protein useful as a therapeutic or prophylactic agent against a  
PT disease caused by an abnormality in a G-protein coupled receptor mediated  
PT signal transduction.  
XX  
PS Claim 1; SEQ ID NO 1; 85pp; English.  
XX  
CC This invention relates to a novel G protein (Gml). The protein is  
CC involved in a G protein-coupled receptor mediated signal transduction.  
CC The protein of the invention has a sequence with a high homology with a  
CC GTP binding site and a GTPase site conserved among G protein alpha  
CC subunits. The protein, the DNA sequence which encodes it and an antibody  
CC specifically recognising the protein of the invention may be useful as a  
CC therapeutic or prophylactic agent against a disease caused by an  
CC abnormality in a G-protein coupled receptor mediated signal transduction.

CC The invention may also be useful for screening for a substance capable of  
 CC regulating a signal transduction mediated by a G-protein-coupled receptor  
 CC and a protein.

XX SQ Sequence 458 AA;

Query Match 100.0%; Score 2400; DB 8; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-200;  
 Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLCYSLRPLFLFGPGDDPCAASEPPVEDAQAPALAPVRAAARDTARTLLPRGEGS 60  
 DB 1 MGLCYSLRPLFLFGPGDDPCAASEPPVEDAQAPALAPVRAAARDTARTLLPRGEGS 60

QY 61 PACARPKADPKPKRQTEQLSAEEREAKEREAVKESRGIDRMRLDQKRDLDQTH 120  
 DB 61 PACARPKADPKPKRQTEQLSAEEREAKEREAVKESRGIDRMRLDQKRDLDQTH 120

QY 121 RLLLLGAGESGKSTIVKQMRILHVGNFNPEKKQKILDIRKNVKDAIVTVSAMSTIIPP 180  
 DB 121 RLLLLGAGESGKSTIVKQMRILHVGNFNPEKKQKILDIRKNVKDAIVTVSAMSTIIPP 180

QY 181 VPLANPENQFRSDYIKSIAPITDFEYSQEFDFHVKKLWDDGVKACFERSNEYQLIDCAQ 240  
 DB 181 VPLANPENQFRSDYIKSIAPITDFEYSQEFDFHVKKLWDDGVKACFERSNEYQLIDCAQ 240

QY 241 YFLERIDSVSLVDYPTDQDLRLCRVLTSGIFETRFQVDKVNFMFVGGQDERRKWIQ 300  
 DB 241 YFLERIDSVSLVDYPTDQDLRLCRVLTSGIFETRFQVDKVNFMFVGGQDERRKWIQ 300

QY 301 CFNDVTAIYVAACSSYNMVIREDNNTNRLRESLDLFESINNNRWLRTISIIILFNKQDM 360  
 DB 301 CFNDVTAIYVAACSSYNMVIREDNNTNRLRESLDLFESINNNRWLRTISIIILFNKQDM 360

QY 361 LAEKVLGKSKI EDPYEPYANYTPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK 420  
 DB 361 LAEKVLGKSKI EDPYEPYANYTPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK 420

QY 421 HCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQYELL 458  
 DB 421 HCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQYELL 458

## RESULT 2

AEAL7292 standard; protein; 458 AA.

XX ID

XX AC AEAL7292;

XX DT 28-JUL-2005 (first entry)

XX DE Human XLGolf protein Seq 2.

XX KW XLGolf; G protein coupled receptor; schizophrenia; psychiatric disorder;

XX KW neuroleptic; gene therapy.

XX OS Homo sapiens.

XX PN WO2005047318-A1.

XX PD 26-MAY-2005.

XX PF 11-NOV-2004; 2004WO-GB004749.

XX PR 11-NOV-2003; 2003US-0519190P.

XX PR 03-SEP-2004; 2004US-0607010P.

XX PA (ASTR ) ASTRAZENECA AB.

XX PA (ASTR ) ASTRAZENECA UK LTD.

XX PI Boetwick RJ, Corradi J, Defay T, Furlong S, Hirata LT, Ravyn V;

XX PI Robbins A;

XX XX

DR WPI; 2005-386333/39.  
 DR N-PSDB; AEAL7291, AEAL7302.

PT New isolated nucleic acid molecule encoding Golf G proteins, useful for  
 PT identifying modulators of G protein coupled receptor activity, or for  
 PT diagnosing or treating schizophrenia and other psychiatric disorders.

PS Claim 8; SEQ ID NO 2; 235pp; English.

XX This invention relates to a novel transcriptional variant of the human  
 CC GNAL gene that encodes a novel splice variant of the G protein alpha  
 CC subunit protein Golf, referred to herein as XLGolf, as well as methods  
 CC for identifying modulators of G protein coupled receptors (GPCRs). Note  
 CC that the N-terminus of the XLGolf protein is altered compared to Golf  
 CC with a different exon 1. Specifically, it refers to contacting the GPCR  
 CC with a test compound, and determining GPCR activity, where a change in  
 CC activity indicates that the compound is a modulator thereof. The present  
 CC invention describes the GPCR as a Gs coupled GPCR that is selected from  
 CC dopamine receptor D1, adenosine A2a receptor, and adrenergic beta-2  
 CC receptor. Accordingly, the composition and methods are useful for  
 CC identifying modulators of GPCR activity, as well as for diagnosing or  
 CC treating schizophrenia and other psychiatric disorders. Furthermore, the  
 CC pharmaceutical compositions derived thereof exhibit neuroleptic activity  
 CC and can be used for gene therapy purposes. This polypeptide sequence is  
 CC the human XLGolf protein of the invention.

XX SQ Sequence 458 AA;

Query Match 100.0%; Score 2400; DB 9; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-200;  
 Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLCYSLRPLFLFGPGDDPCAASEPPVEDAQAPALAPVRAAARDTARTLLPRGEGS 60  
 DB 1 MGLCYSLRPLFLFGPGDDPCAASEPPVEDAQAPALAPVRAAARDTARTLLPRGEGS 60

QY 61 PACARPKADPKPKRQTEQLSAEEREAKEREAVKESRGIDRMRLDQKRDLDQTH 120  
 DB 61 PACARPKADPKPKRQTEQLSAEEREAKEREAVKESRGIDRMRLDQKRDLDQTH 120

QY 121 RLLLLGAGESGKSTIVKQMRILHVGNFNPEKKQKILDIRKNVKDAIVTVSAMSTIIPP 180  
 DB 121 RLLLLGAGESGKSTIVKQMRILHVGNFNPEKKQKILDIRKNVKDAIVTVSAMSTIIPP 180

QY 181 VPLANPENQFRSDYIKSIAPITDFEYSQEFDFHVKKLWDDGVKACFERSNEYQLIDCAQ 240  
 DB 181 VPLANPENQFRSDYIKSIAPITDFEYSQEFDFHVKKLWDDGVKACFERSNEYQLIDCAQ 240

QY 241 YFLERIDSVSLVDYPTDQDLRLCRVLTSGIFETRFQVDKVNFMFVGGQDERRKWIQ 300  
 DB 241 YFLERIDSVSLVDYPTDQDLRLCRVLTSGIFETRFQVDKVNFMFVGGQDERRKWIQ 300

QY 301 CFNDVTAIYVAACSSYNMVIREDNNTNRLRESLDLFESINNNRWLRTISIIILFNKQDM 360  
 DB 301 CFNDVTAIYVAACSSYNMVIREDNNTNRLRESLDLFESINNNRWLRTISIIILFNKQDM 360

QY 361 LAEKVLGKSKIEDYEPYANYTPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK 420  
 DB 361 LAEKVLGKSKIEDYEPYANYTPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK 420

QY 421 HCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQYELL 458  
 DB 421 HCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQYELL 458

## RESULT 3

ADG74746

ID ADG74746 standard; protein; 448 AA.

XX AC ADG74746;

XX DT 22-APR-2004 (first entry)

XX XX

DE Mouse G-protein Gml amino acid sequence.

XX G protein; Gml; G protein-coupled receptor mediated signal transduction; 420

KW GTP binding site; GTPase site; G protein alpha subunit; 410

KW signal transduction; G-protein-coupled receptor; mouse; murine.

XX Mus musculus.

XX EPI382613-A1.

XX 21-JAN-2004.

XX 09-JUL-2003; 2003EP-00015519.

XX 16-JUL-2002; 2002JP-00206841.

XX 19-DEC-2002; 2002JP-00367778.

XX 31-MAR-2003; 2003JP-00095955.

XX (SUMO ) SUMITOMO CHEM CO LTD.

XX Takahashi Y, Matsumoto Y, Oeda K;

XX WPI; 2004-111483/12.

XX N-PSDB; ADG74748.

XX New protein useful as a therapeutic or prophylactic agent against a

PT disease caused by an abnormality in a G-protein coupled receptor mediated

PT signal transduction.

XX Claim 1; SEQ ID NO 25; 85pp; English.

XX This invention relates to a novel G protein (Gml). The protein is

CC involved in a G protein-coupled receptor mediated signal transduction.

CC The protein of the invention has a sequence with a high homology with a

CC GTP binding site and a GTPase site conserved among G protein alpha

CC subunits. The protein, the DNA sequence which encodes it and an antibody

CC specifically recognising the protein of the invention may be useful as a

CC therapeutic or prophylactic agent against a disease caused by an

CC abnormality in a G-protein coupled receptor mediated signal transduction.

CC The invention may also be useful for screening for a substance capable of

CC regulating a signal transduction mediated by a G-protein-coupled receptor

CC and a protein. The present sequence is that of the mouse Gml protein

CC which is related to the human Gml protein of the invention.

XX SQ Sequence 448 AA;

Query Match 88.5%; Score 2124; DB 8; Length 448;

Best Local Similarity 91.0%; Pred. No. 6.7e-176; Indels 10; Gaps 3;

Matches 417; Conservative 6; Mismatches 25; Indels 10; Gaps 3;

QY 1 MGLCYSLRPLLFSGGDDPCAAEPVVEDAQAPAPALAPVRAAARDTARTLLPRGGE 60

DB 1 MGLCYSLRPLLFSGPETPCNASEPCDAQPSAPAPASIPAPA--PVGTLRRGGRI 58

QY 61 PACARPKADPKPKRQTEQLSABEREAEREAVKREAVKRGIDRLMRDQKRDLOQTH 120

DB 59 VANARPPGE--LQSRREQQLRABEREA-----KEARKVSRGIDRLMRDQKRDLOQTH 110

QY 121 RLLLLGSGSKSTIVKQMLHNGFNPEKKQILDIRKVKDAITVVSAMSTIIPP 180

DB 111 RLLLLGSGSKSTIVKQMLHNGFNPEKKQILDIRKVKDAITVVSAMSTIIPP 170

QY 181 VPLANPENQFRSDYIKSIAPITDFEYSOEFDHVKLWDDGKACFERSNEYQLIDCAQ 240

DB 171 VPLANPENQFRSDYIKSIAPITDFEYSOEFDHVKLWDDGKACFERSNEYQLIDCAQ 230

QY 241 YFLERIDSVSLVDYPTDQDLRLCRVLTSGIFETFRFQVDKVNFMFVGGQDERRKWIQ 300

DB 231 YFLERIDSVSLVDYPTDQDLRLCRVLTSGIFETFRFQVDKVNFMFVGGQDERRKWIQ 290

QY 301 CFNDVTAIIVAAACSSYNNVIRENNTNRLRESLDLFESINNNRLWTISIIILFNKQDM 360

DB 291 CFNDVTAIIVAAACSSYNNVIRENNTNRLRESLDLFESINNNRLWTISIIILFNKQDM 350

QY 361 LAEKVLAKGSKIEDYFPEYANYTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK 420

DB 351 LAEKVLAKGSKIEDYFPEYANYTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK 410

QY 421 HVCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQYELL 458

DB 411 HVCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQYELL 448

RESULT 4

ADG74747

ID ADG74747 standard; protein; 450 AA.

XX ADG74747;

AC ADG74747;

XX 22-APR-2004 (first entry)

XX Rat G-protein Gml amino acid sequence.

XX G protein; Gml; G protein-coupled receptor mediated signal transduction;

KW GTP binding site; GTPase site; G protein alpha subunit;

KW signal transduction; G-protein-coupled receptor; rat.

XX Rattus norvegicus.

XX EPI382613-A1.

XX 21-JAN-2004.

XX 09-JUL-2003; 2003EP-00015519.

XX 16-JUL-2002; 2002JP-00206841.

XX 19-DEC-2002; 2002JP-00367778.

XX 31-MAR-2003; 2003JP-00095955.

XX (SUMO ) SUMITOMO CHEM CO LTD.

XX Takahashi Y, Matsumoto Y, Oeda K;

XX WPI; 2004-111483/12.

XX N-PSDB; ADG74749.

XX New protein useful as a therapeutic or prophylactic agent against a

PT disease caused by an abnormality in a G-protein coupled receptor mediated

PT signal transduction.

XX Claim 1; SEQ ID NO 26; 85pp; English.

XX This invention relates to a novel G protein (Gml). The protein is

CC involved in a G protein-coupled receptor mediated signal transduction.

CC The protein of the invention has a sequence with a high homology with a

CC GTP binding site and a GTPase site conserved among G protein alpha

CC subunits. The protein, the DNA sequence which encodes it and an antibody

CC specifically recognising the protein of the invention may be useful as a

CC therapeutic or prophylactic agent against a disease caused by an

CC abnormality in a G-protein coupled receptor mediated signal transduction.

CC The invention may also be useful for screening for a substance capable of

CC regulating a signal transduction mediated by a G-protein-coupled receptor

CC and a protein. The present sequence is that of the rat Gml protein which

CC is related to the human Gml protein of the invention.

XX SQ Sequence 450 AA;

Query Match 88.0%; Score 2113; DB 8; Length 450;

Best Local Similarity 90.3%; Pred. No. 6.1e-175;

Matches 415; Conservative 7; Mismatches 26; Indels 12; Gaps 4;

QY 1 MGLCYSLRPLLFSGGDDPCAAEPVVEDAQAPAPALAPVRAAARDTARTLLPRGGE 58

DB 1 MGLCYSLRPLLFSGGDDPCEDSEPCAEQPSAAPAPAPAPAPA--PVGTLRRGGD 58

QY 59 GSPACARPKADPKPKRQTEQLSABEREAEREAVKREAVKRGIDRLMRDQKRDLOQ 118

Db 59 RIPASRSPVE--LQNRROEQLRABEREAA-----KEARKVSRGIDRMLRQKRDLLQ 110  
Qy 119 THRLLLGAGESGKSTIVKQMRILHVNQFNPBEKKQKILDIRKNVKDAIVTIVSAMSTII 178  
Db 111 THRLLLGAGESGKSTIVKQMRILHVNQFNPBEKKQKILDIRKNVKDALVTIISAMSTII 170  
Qy 179 PPVPLANPENQFSDYIKTSIAPTIDREYSOEFFPDHVKKLDWDDGKACFERSNEYQIDC 238  
Db 171 PPVPLANPENQFSDYIKTSIAPTIDREYSOEFFPDHVKKLDWDDGKACFERSNEYQIDC 230  
Qy 239 AQYFLERIDSVLDVYPTDQDLLRCRVLTSGIFETRFQVDKYNFMDVGGQDERRKW 298  
Db 231 AQYFLERIDSVLDVYPTDQDLLRCRVLTSGIFETRFQVDKYNFMDVGGQDERRKW 290  
Qy 299 IQCFNDVTALIIYVAACSSYNMVRDNTNRLRESLDLFSIWNRLRTISIIILFNKQ 358  
Db 291 IQCFNDVTALIIYVAACSSYNMVRDNTNRLRESLDLFSIWNRLRTISIIILFNKQ 350  
Qy 359 DMLAEKVLACKSKIEDYFPEYANYTVPEDATPDAGDPKVTRAKFFIRDLFLRISTATGD 418  
Db 351 DMLAEKVLACKSKIEDYFPEYANYTVPEDATPDAGDPKVTRAKFFIRDLFLRISTATGD 410  
Qy 419 GKHYCPHFTCAVDTENIRRVFNDCRDIIQRMHLKQYELL 458  
Db 411 GKHYCPHFTCAVDTENIRRVFNDCRDIIQRMHLKQYELL 450

RESULT 5  
ABB09272  
ID ABB09272 standard; protein; 381 AA.  
XX  
AC ABB09272;  
XX  
DT 10-JUL-2002 (first entry)  
XX  
DE G protein-coupled receptor (GPCR) >g-olf SEQ ID NO:18.  
XX  
KW Target activated nucleic acid biosensor; signalling moiety; GPCR;  
KW nucleic acid sensor; detection; engineering; drug optimisation;  
KW G protein-coupled receptor.  
XX  
OS Homo sapiens.  
XX  
PN WO200222882-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 13-SEP-2001; 2001WO-US028835.  
XX  
PR 13-SEP-2000; 2000US-0232454P.  
XX  
PA (ARCH-) ARCHEMIX CORP.  
XX  
PI Stanton M, Epstein D, Hamaguchi N;  
XX  
DR WPI; 2002-393977/42.  
XX  
PT Nucleic acid sensor for detecting target molecule, comprises target  
PT molecule activation site and optical signalling unit that changes its  
PT optical properties upon allosteric modulation sensor after recognition of  
PT target.  
XX  
PS Example 12; Page 88-89; 144pp; English.  
XX  
CC The present invention describes a nucleic acid sensor molecule (I)  
CC comprising a target molecule activation site comprising a structure that  
CC recognises a target molecule and an optical signalling unit including at  
CC least one nucleotide coupled to a signalling moiety that changes its  
CC optical properties upon allosteric modulation of (I) following  
CC recognition of the target molecule. (I) is useful for detecting a target  
CC molecule associated with a pathological condition or genetic alteration.  
CC (I) is useful for identifying a drug compound, by identifying a nucleic

CC acid biosensor-based molecule profile of target molecules associated with  
CC a disease trait in a patient, administering a candidate compound to the  
CC patient, and monitoring changes in the profile. Alternately, the method  
CC involves identifying a number of pathway target molecules, administering  
CC a candidate compound to a patient having a disease trait, and monitoring  
CC changes in the structure, level or activity of two or more of the pathway  
CC target molecules using (I). The profile of target molecules or the  
CC changes in the structure is compared to the profile of a reference  
CC healthy or diseased population. (I) is useful in multiple assays, for the  
CC detection of target molecule. (I) is also useful in diagnostic  
CC applications and drug optimisation. The present sequence represents a G  
CC protein-coupled receptor, which is used in an example from the present  
CC invention  
XX  
SQ Sequence 381 AA;  
Query Match 75.8%; Score 1819; DB 5; Length 381;  
Best Local Similarity 95.0%; Pred. No. 1.9e-149;  
Matches 345; Conservative 9; Mismatches 9; Indels 0; Gaps 0;  
Qy 96 KEARKVSRGIDRMLRQKRDLLQTHRLLLGAGESGKSTIVKQMRILHVNQFNPBEKKQK 155  
Db 19 KERREANKKIEKQLQKRLAYKATHRLLLGAGESGKSTIVKQMRILHVNQFNPBEKKQK 78  
Qy 156 ILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFSDYIKTSIAPTIDREYSOEFFPDHVK 215  
Db 79 ILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFSDYIKTSIAPTIDREYSOEFFPDHVK 138  
Qy 216 KLWDDGKACFERSNEYQIDCAQYFLERIDSVLDVYPTDQDLLRCRVLTSGIFETR 275  
Db 139 KLWDDGKACFERSNEYQIDCAQYFLERIDSVLDVYPTDQDLLRCRVLTSGIFETR 198  
Qy 276 FQVDKYNFMDVGGQDERRKWIQCFNDVTALIIYVAACSSYNMVRDNTNRLRESLD 335  
Db 199 FQVDKYNFMDVGGQDERRKWIQCFNDVTALIIYVAACSSYNMVRDNTNRLRESLD 258  
Qy 336 LFESIWNRLRTISIIILFNKQDMLAEKVLACKSKIEDYFPEYANYTVPEDATPDAGSD 395  
Db 259 LFESIWNRLRTISIIILFNKQDMLAEKVLACKSKIEDYFPEYANYTVPEDATPDAGSD 318  
Qy 396 PKVTRAKFFIRDLFLRISTATGDGKHVCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQY 455  
Db 319 PKVTRAKFFIRDLFLRISTATGDGKHVCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQY 378  
Qy 456 ELL 458  
Db 379 ELL 381  
RESULT 6  
ADC09607  
ID ADC09607 standard; protein; 381 AA.  
XX  
AC ADC09607;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human G-protein coupled receptor-related protein, SEQ ID 18.  
XX  
KW Nucleic acid sensor molecule; ligase; cis-hammerhead; protein kinase;  
KW human; G-protein coupled receptor.  
XX  
OS Homo sapiens.  
XX  
PN WO2003014375-A2.  
XX  
PD 20-FEB-2003.  
XX  
PF 09-AUG-2002; 2002WO-US025319.  
XX  
PR 09-AUG-2001; 2001US-0311378P.  
XX  
PR 21-AUG-2001; 2001US-0313932P.  
XX  
PR 13-SEP-2001; 2001US-00952680.



PR 13-NOV-2001; 2001US-0338186P.  
PR 18-JAN-2002; 2002US-0349959P.  
PR 13-MAR-2002; 2002US-0364486P.  
PR 25-MAR-2002; 2002US-0367991P.  
PR 04-APR-2002; 2002US-0369887P.  
PR 01-MAY-2002; 2002US-0376744P.  
PR 31-MAY-2002; 2002US-0385097P.  
XX (ARCH-) ARCHEMIX CORP.  
XX Stanton M, Epstein D, Hamaguchi N, Kurz M, Keefe T, Wilson C;  
PI Grate D, Marshall KA, Mccauley T, Kurz J;  
XX WPI; 2003-300534/29.  
XX Nucleic acid sensor molecule, for identifying/detecting protein kinase in  
PT a sample, comprises a target modulation domain which recognizes a target  
PT molecule, a linker domain, a catalytic domain, and an optical signal  
PT generator.  
XX  
XX Example 5; SEQ ID NO 18; 423pp; English.  
XX  
CC The present invention relates to nucleic acid sensor molecules (I), which  
CC comprise a target modulation domain that recognizes a target molecule  
CC (TM), a linker domain, a catalytic domain, and an optical signal  
CC generating unit. The catalytic domain comprises a ligase or cis-  
CC hammerhead. (I) are useful for identifying or detecting TM in a sample,  
CC preferably a protein kinase in a sample. Target molecules include  
CC proteins, post-translationally modified forms of proteins, peptides,  
CC nucleic acids, oligosaccharides, nucleotides, metabolites, drugs, toxins,  
CC biohazards, ions, carbohydrates, polysaccharides, hormones, receptors,  
CC antigens, antibodies, viruses, metabolites, co-factors, drugs, dyes,  
CC nutrients, growth factors, cAMP, cAMP or cGMP, protein kinase,  
CC phosphorylated protein kinase, extracellular signal regulated kinase  
CC (ERK), a component or product of mitogen activated protein (MAP) kinase  
CC pathway, a MAP kinase pathway associated protein, an extracellular  
CC component of MAP kinase pathway, a component of ERK1/2 MAP, JNK MAP or  
CC P38 MAP kinase pathway, an endogenous form of MAP kinase (MEK), MAP  
CC kinase kinase, or MAP kinase (MEKK), or RAF kinase, Ras protein,  
CC phosphatase, GTP binding protein, G-protein coupled receptor (GPCR),  
CC cytokine, growth factor, cellular metabolite, small molecule or lysosome.  
CC (I) are also useful for identifying a modulator of protein kinase  
CC activity. In an example from the invention, nucleic acid sensor molecules  
CC which signal human G-protein coupled receptors e.g. the present sequence,  
CC were obtained.  
XX  
SQ Sequence 381 AA;  
Query Match 75.8%; Score 1819; DB 7; Length 381;  
Best Local Similarity 95.0%; Pred. No. 1.9e-149;  
Matches 345; Conservative 9; Mismatches 9; Indels 0; Gaps 0;  
QY 96 KEARKVSRGIDRMRLDOKRDLQOQTHLLGAGSGKSTIVKQWRIHVGNGFPEKKQK 155  
DB 19 KERREANKKEKQKQERLAYKATHRLLLGAGSGKSTIVKQWRIHVGNGFPEKKQK 78  
QY 156 ILDIRKNVKDAIVTIVSAMSTIIPVPLANPENFRSDYIKSIAPITDFYISOEFFDHVK 215  
DB 79 ILDIRKNVKDAIVTIVSAMSTIIPVPLANPENFRSDYIKSIAPITDFYISOEFFDHVK 138  
QY 216 KLWDEGVKACFRSNEYQLIDCAQYFLERIDSLSVLDYVPTDQDLRCRVLTSIGIPETR 275  
DB 139 KLWDEGVKACFRSNEYQLIDCAQYFLERIDSLSVLDYVPTDQDLRCRVLTSIGIPETR 198  
QY 276 FOVDKVNFMFDVGGQDERKWTQCFNDVTAILIYVAACSSYNNVIBEDNNTNRLRESLD 335  
DB 199 FOVDKVNFMFDVGGQDERKWTQCFNDVTAILIYVAACSSYNNVIBEDNNTNRLRESLD 258  
QY 336 LFSIWNRRMLRTSIIILFNKQDLAEKVLAGSKIEDYFPPEYANTVTPEDATPDAGED 395  
DB 259 LFSIWNRRMLRTSIIILFNKQDLAEKVLAGSKIEDYFPPEYANTVTPEDATPDAGED 318  
QY 396 PKVTRAKFFIRDLFLRISTATGDKGKCYPHFTCAVDNTNIRRVFNDCRDIIQRMHLKQY 455

Db 319 PKVTRAKFFIRDLFLRISTATGDKGKCYPHFTCAVDNTNIRRVFNDCRDIIQRMHLKQY 378  
QY 456 ELL 458  
Db 379 ELL 381  
RESULT 7  
ADE61907  
ID ADE61907 standard; protein; 381 AA.  
XX  
AC ADE61907;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human Protein P38405, SEQ ID NO 7836.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
PN W02003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEHO ) GEN HOSPITAL CORP.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
XX WPI; 2003-268312/26.  
DR GENBANK; P38405.  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.  
XX  
CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at

[illegible]





```

PR XX 26-NOV-2001; 2001US-0333347P.
XX PA (GEO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX PI Woelf C, D'urao D, Befort K, Costigan M;
XX DR WPI; 2003-368912/26.
XX DR GENBANK; P38406.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX CC The invention discloses a composition comprising two or more isolated rat
CC CC or human polynucleotides or a polynucleotide which represents a fragment,
CC CC derivative or allelic variation of the nucleic acid sequence. Also
CC CC claimed are a vector comprising the novel polynucleotide, a host cell
CC CC comprising the vector, a method for identifying a nucleotide sequence
CC CC which is differentially regulated in an animal subjected to pain and a
CC CC kit to perform the method, an array, a method for identifying an agent
CC CC that increases or decreases the expression of the polynucleotide sequence
CC CC that is differentially expressed in neuronal tissue of a first animal
CC CC subjected to pain, a method for identifying a compound which regulates
CC CC the expression of a polynucleotide sequence which is differentially
CC CC expressed in an animal subjected to pain, a method for identifying a
CC CC compound that regulates the activity of one or more of the
CC CC polynucleotides, a method for producing a pharmaceutical composition, a
CC CC method for identifying a compound or small molecule that regulates the
CC CC activity in an animal of one or more of the polypeptides given in the
CC CC specification, a method for identifying a compound useful in treating
CC CC pain and a pharmaceutical composition comprising the one or more
CC CC polypeptides or their antibodies. The polynucleotide or the compound that
CC CC modulates its activity is useful for preparing a medicament for treating
CC CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC CC the specification) which is differentially expressed during pain. Note:
CC CC The sequence data for this patent did not form part of the printed
CC CC specification, but was obtained in electronic form directly from WIPO at
CC CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 381 AA;

Query Match      .   75.5%; Score 1811; DB 7; Length 381;
Best Local Similarity 94.8%; Pred. No. 9.7e-149;
Matches 344; Conservative 9; Mismatches 10; Indels 0; Gaps 0

QY    96 KEARKVSIGIDRMLDQRDQLQOTHRLLLLLGAGSGSKSTIVKQWRIILHVGNFPEEKOK 155
     ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB    19 KERREANKKIELOKERLAVKATHRLLLLGAGSGSKSTIVKQWRIILHVGNSPEEKOK 78

QY    156 ILDIRKNVKDAITVTIVSAMSTIIPVPPLANPENOFRSDYIKTSAPITDPYSEPPDHVK 215
     ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB    79 ILDIRKNVKDAITVTIVSAMSTIIPVPPLANPENOFRSDYIKTSAPITDPYSEPPDHVK 138

QY    216 KLWDDEGVKACERSNEYOLIDCAQYFLERIDSLSVDYTPTDDLLRCRVLTSGIFETR 275
     ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB    139 KLWDDEGVKACERSNEYOLIDCAQYFLERIDSLSVDYTPTDDLLRCRVLTSGIFETR 198

QY    276 FOVDKVNFMHDVGQDRERRKWTCFNVDVTAIIVAACSSYNMVRDNNTNRLRESID 335
     ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB    199 FOVDKVNFMHDVGQDRERRKWTCFNVDVTAIIVAACSSYNMVRDNNTNRLRESID 258

QY    336 LPFESINNRWLRTISIIILFLNKQDMLEAKVLGASKIEDYPPEYANTVPEDATPDAGD 395
     ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB    259 LPFESINNRWLRTISIIILFLNKQDMLEAKVLGASKIEDYPPEYANTVPEDATPDAGD 318

QY    396 PKVTRAKFFIRDLFURISTATGDGHGYCPHFTCAVDTENIRRVFNDCRDIIORMHLKOY 455
     ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB    319 PKVTRAKFFIRDLFURISTATGDGHGYCPHFTCAVDTENIRRVFNDCRDIIORMHLKOY 378

```

Qy	59	GSAPCARPKADXP-KEKQRTQESAEBREAAKREAVKEARKVSGRIGDRMLDRDQKRDQD IL17	
Db	355	GLSECTRSRSLSPGKAKDPMERRRKQMKREIAENREQRADKSKSLIDKQLEEBEKDYM 414	
Qy	118	QTHRLLLLGAGSGSKSTIVKQNRILHVNFGNPE-EKKQKILDIRKNVKNDAIVTIVSAMST 176	
Db	415	CTHRLLLGAGSGSKSTIVKQNRILHVNFGNDSEKATKVODIKNNLKEAETIVAAASN 474	
Qy	177	IIPVPPLANPENQFRSDYIKSIAPITDPEYSQEFDFHWKKLWDDGKVACFERSNEYQLI 236	
Db	475	LVPVVELANPENQFRVDYILSVNVPNDFPFPEYEHAKALWEDGVRACYSERSNEYQLI 534	
Qy	237	DCAQYFLERIDSVLVDYPTDQDLLRCRVLTSGIFETRFOVDKFNPHMFDVGQDERR 296	
Db	535	DCAQYFLDIDVIKQADYVPSQDILLCRVLTSGIFETKFQVDKFNPHMFDVGQDERR 594	
Qy	297	KWIOCFENDVTALIIYVAACSSNMVITRENNNTNRLESIDLPESTWNNRWLRTIISILFLN 356	
Db	595	KWIOCFENDVTALIIYVASSSNMVITRENNQTRNQLQEALNFKSIWNNRWLRTISVILFLN 654	
Qy	357	KQDMLAEKVLGAKSKIEDYFPFAYANTVPEDATPDAGEDPKVTRAKFPFIRDLFLRISTAT 416	
Db	655	KQDMLAEKVLGAKSKIEDYFPFARYTTPEDATPEPGEDPRVTRAKFYIRDEFRLISTAS 714	
Qy	417	GDGKHVCYPHFTCAVDTENIRRVFNCDCRDIITQRMHLQOYELL 458	
Db	715	GDGRHYCYPHFTCAVDTENIRRVFNDCRDIITQRMHLQOYELL 756	

RESULT 13	
ABP97657	
ID	ABP97657 standard; protein; 756 AA.
XX	
AC	ABP97657;
XX	
DT	16-MAY-2003 (first entry)
XX	
DE	Amino acid sequence of a murine GNAS polypeptide G protein xi-alpha.
XX	
KW	GNAS; lymphoma; drug screening; leukemia; wound; inflammation;
KW	immunotherapy; oncogene; vaccine; G protein xi-alpha.
KW	

XX	OS	Mus musculus.	
XX	OS		
XX	PN	WO2003006688-A2.	
XX	XX		
XX	PD	23-JAN-2003.	
XX	XX		
XX	PF	15-JUL-2002; 2002WO-EP007853.	
XX	XX		
XX	PR	13-JUL-2001; 2001US-00905390.	
XX	PR	24-SEP-2001; 2001WO-US029798.	
XX	XX		
XX	PA	(UYAA-) UNIV AARHUS.	
XX	XX		
XX	PI	Pedersen FS, Sorensen AB, Hernandez JM;	
XX	DR	WPI; 2003-229494/22.	
XX	DR	N-PSDB; ABZ68699.	
XX	XX		
XX	PT	Screening drug candidates which can modulate Gnas proteins for treating	
XX	PT	lymphoma or leukemia, comprises contacting a cell expressing a GNAS gene	
XX	PT	with a candidate drug and determining the effect of the drug on gene	
XX	PT	expression.	
XX	XX		
XX	PS	Disclosure; Page 49; 79pp; English.	
XX	XX		
XX	CC	The present sequence represents the GNAS polypeptide G protein XI-alpha.	
XX	CC	GNAS is a complex locus encoding multiple proteins, and GNAS genes are	
XX	CC	implicated in lymphoma and leukemias. The specification describes a	
XX	CC	method for screening drug candidates. The method comprises providing a	
XX	CC	cell expressing a GNAS gene, adding a drug candidate to the cell, and	
XX	CC	determining the effect of the drug candidate on the expression of a GNAS	

	CC	gene. The method is useful for screening drugs which can modulate GNAS proteins and which can be used to treat lymphomas, leukemia, wounds and inflammation. Antibodies raised against a GNAS protein are useful in passive or active immunotherapy, for inhibiting oncogenes, and for diagnosing lymphoma from blood samples. GNAS proteing may be used as markers of lymphoma, in in situ imaging techniques, in drug screening assays or by evaluating the effect of drug candidates on a gene or protein expression profile, and in generating antibodies against the protein. GNAS genes may be used as probes to determine the number of copies of the GNAS gene in the genome or to determine the chromosomal location of the GNAS gene, and as DNA vaccines
XX	CC	
XX	CC	
SQ	Sequence 756 AA;	
	Query Match	65.0%; Score 1559; DB 6; Length 756;
	Best Local Similarity	72.6%; Pred. No. 2.2e-126;
	Matches 292; Conservative	48; Mismatches 60; Indels 2; Gaps 2;
Qy	59	GSPACARPKADKP-KEKQRQTQSLSAEREAKREAVKEARKVSRGIDRMRLDKOKRDILQ 117     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   : Db 355 GLSECTRSRSLSPGKAPDPMEEERRKKMRKEAIEMREQKRADKGSKLJDKOLEBEKMAYM 414     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy	118	OQHRIILLGGESGSTIVKOMRTLLHVNNGNPE-EKKOKILDIRKNVKDAIWTIVSAMST 176     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Dd	415	CTRHLIIIGAGESGSKTIVQMRTLHVNFSGDSKATKVODIKNNLKEARLETIVAAMSN 474     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy	177	IIPVPPLANPENQRSPIYKIAPTITDFEYSQEFPFDHKKLWDEGVKACFERSNEYQLI 236     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Dd	475	LVPVELANPENQFRVDYLISVMNVPNFDPFEFYEHAKALWEDEGVRACYERSNEYQLI 534     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy	237	DCAQYFLERIDSVLSDYTPTDQDLRCRVLTSGIFETRPFOVDKVFHFMDVGGOQDRRR 296     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Dd	535	DCAQYFLKDVIKAQDVPSDQDLLRCRVLTSGIFTFKQVDKVFHFMDVGGOQDRRR 594     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy	297	KWIQCNDVTAILIYAACSSYNMVIRDNTNRLESIDLPESTWNNEWLRTSILFILFN 356     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Dd	595	KWIQCNDVTAILIYFAVVASSYNMVIREDNQNRIQAELNLFPSIWNNEWLRTSILFILFN 654     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy	357	KQDMIAEKVLAGSKIEDIYPEYANYTVPEDATPDAGEDPVKTKRAFFIRDLFLRAISTAT 416     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Dd	655	KQDMAEKVLAGSKIEDIYPARYTTPEDATEPGEDEVPRVTRAIFYRDEFLEASTAS 714     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy	417	GDGRHYCYPHFTCAVDENIRRVENDCRDIIRMHLKOYEELL 458     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Dd	715	GDGRHYCYPHFTCAVDENIRRVENDCRDIIRMHLROYELL 756     : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
RESULT 14		
ADQ26060	ID	
ADQ26060 standard; protein; 909 AA.	ADQ26060	
XX AC ADQ26060;		
XX DT		
XX TT 23-SEP-2004 (first entry)		
XX DE Guanine nucleotide binding protein, alpha stimulating complex locus #2.		
XX KW BCBC; brain capillary endothelial cell; astrocyte; permeability;		
KW endothelial cell; pro-barrier; PB;		
KW microvascular permeability-modifying disorder;		
KW neurodegenerative disorder; cerebrovascular accident;		
KW Alzheimer's disease; vascular-related dementia;		
KW Creutzfeldt-Jacob disease; bovine spongiform encephalopathy;		
KW parkinson's disease; brain trauma; multiple sclerosis;		
KW peripheral disorder; septic shock; hepatic encephalopathy;		
KW diabetic hypertension; diabetic microangiopathy; sleeping sickness;		
KW neuropsychiatric disorder; depression; autism; schizophrenia; psychosis;		
KW CNS disorder; brain tumor; epilepsy; migraine; narcolepsy; insomnia;		
KW chronic fatigue syndrome; mountain sickness; AIDS-related dementia;		
KW angiogenesis-related disorder; proliferative vitreoretinopathy;		
KW rheumatoid arthritis; Crohn's disease; atherosclerosis;		
KW ovarian hyperstimulation; psoriasis; endometriosis; neovascularisation;		
KW restenosis; balloon angioplasty; scar tissue overproduction		



peripheral vascular disease; hypertension; inflammatory vasculitides; Reynaud's disease; Reynaud's phenomenon; aneurysm; arterial restenosis; thromboangiitis; lymphangitis; lymphedema; wound healing; tissue repair; ischaemia reperfusion injury; angina; myocardial infarction; chronic heart condition; osteoporosis; Guanine nucleotide binding protein; alpha stimulating complex locus; GNAS.

XX Homo sapiens.

XX WO2004056386-A2.

XX 08-JUL-2004.

XX 19-DEC-2003; 2003WO-NL000915.

XX 19-DEC-2002; 2002EP-00080503.

XX 25-APR-2003; 2003US-0465234P.

XX (UYLE-) RIJKSUNIV LEIDEN.

XX Gaillard PJ, De Boer AG, Brink A;

XX WPI; 2004-500264/47.

XX Modulating the permeability of endothelial cells, useful for treating or preventing e.g. neurodegenerative disorders, comprises altering in the endothelial cells the activity or the steady-state level of a pro-barrier polypeptide.

XX Claim 1; SEQ ID NO 14; 222pp; English.

CC This sequence represents guanine nucleotide binding protein, alpha stimulating complex locus encoded by the GNAS gene. This protein is upregulated in BCEC (brain capillary endothelial cells) by physical co-culture with astrocytes. This protein sequence may be used in the method of the invention for modulating the permeability of endothelial cells. The method comprises altering in the endothelial cells the activity or the steady-state level of a pro-barrier (PB) polypeptide having an amino acid sequence having at least 90% identity with a claimed amino acid sequence selected from ADQ26047ADQ26048ADQ26049ADQ26050ADQ26051ADQ26052ADQ26053ADQ26054ADQ26055ADQ26056ADQ26057ADQ26058ADQ26059ADQ26060ADQ26061ADQ26062ADQ26063ADQ26064ADQ26065ADQ26066ADQ26067ADQ26068ADQ26069ADQ26070ADQ26071ADQ26072ADQ26073ADQ26074ADQ26075ADQ26076ADQ26077ADQ26078ADQ26079ADQ26080ADQ26081ADQ26082ADQ26083ADQ26084ADQ26085ADQ26086ADQ26087ADQ26088ADQ26089ADQ26090ADQ26091ADQ26092ADQ26093ADQ26094ADQ26095ADQ26096ADQ26097ADQ26098ADQ26099. The PB polypeptides, nucleic acid molecules encoding them, antagonists or a gene therapy vector comprising an antisense nucleotide sequence capable of inhibiting the expression of the nucleotide sequence encoding a PB polypeptide, are useful in the manufacture of a composition for treating or preventing a microvascular permeability-modifying disorder, including neurodegenerative disorders (e.g. cerebrovascular accidents, Alzheimer's disease, vascular-related dementia, Creutzfeldt-Jacob disease, bovine spongiform encephalopathy, Parkinson's disease, brain trauma, multiple sclerosis), peripheral disorders with a CNS component (such as septic shock, hepatic encephalopathy, (diabetic) hypertension, diabetic microangiopathy or sleeping sickness), neuropsychiatric disorders (e.g. depression, autism, schizophrenia and other psychoses; other CNS disorders (brain tumours, epilepsy, migraine, narcolepsy, insomnia), chronic fatigue syndrome, mountain sickness, or AIDS-related dementia); and angiogenesis-related disorders (proliferative atherosclerosis, rheumatoid arthritis, Crohn's disease, associated with neovascularisation, restenosis subsequent to balloon angioplasty, scar tissue overproduction, peripheral vascular disease, hypertension, inflammatory vasculitides, Reynaud's disease, Reynaud's phenomenon, aneurysms, arterial restenosis, thrombophlebitis, lymphangitis, lymphedema, wound healing and tissue repair, ischaemia reperfusion injury, angina, myocardial infarctions, chronic heart conditions, or osteoporosis). The PB polypeptide may also be used in the manufacture of a composition for reversibly increasing the microvascular permeability in a subject.

SQ Sequence 909 AA;

Query Match 65.0%; Score 1559; DB 8; Length 909;

Best Local Similarity 68.6%; Pred. No. 2.9e-126;

Matches 300; Conservative 49; Mismatches 68; Indels 20; Gaps 4;

QY 42 RAAARDATRTLLPR---GGEGSPACARPADKP-KEKROQTEQLSAEEREAAKEREAVKE 97

DB 473 RKQORLLRNFLVQAFGGCFGRSESPQPKASRLKVKVPLAEKRQMRKEALEKRAQKR 532

QY 98 A-RKVSRGIDRMRLDQKRDILQOQTHRLLLGAGSGSKSTIVKQMRILHVNFGNPE----- 150

DB 533 AEKRSKLIDKLODEKMGYMCYHTRLLLLGAGSGSKSTIVKQMRILHVNFGNFGEGGEDP 592

QY 151 -----EKOKILDIRKNKDAIVTIVSAMSTIIPVPLANPENQFSDYIKSTAPI 201

DB 593 QAARSNSDGEKATKVQDIKNLKEALETIVAAASNLVPPVELANPENQFSDYILSVNMV 652

QY 202 TDFEYSQEPFDHVKKLWDDEGVKACFERSNEYOLIDCAQYFLERIDSLSVLTPTDODL 261

DB 653 PDFDFPPEFYEHAKEALWEDEGVACYSNEYOLIDCAQYFLDKIDVIKQADVPSQDL 712

QY 262 LRCRVLTSGIFETRFQVDKVNFMFVDVGGORERRKWIQCFNDVTALIIYVAACSSNMVI 321

DB 713 LRCRVLTSGIFETRFQVDKVNFMFVDVGGORERRKWIQCFNDVTALIIYVAACSSNMVI 772

QY 322 REDNNTNRLRESLDLFPESINNNRWLRTISIIILFNKQDMLAEKVLAKGSKIETYFPEYAN 381

DB 773 REDNNTNRLRESLDLFPESINNNRWLRTISIIILFNKQDMLAEKVLAKGSKIETYFPEYAN 832

QY 382 YTVPEATPDAGDDPKVTRAKFIRDLFLRISTATGDKHYCYPHTCVADTENNIRRVFN 441

DB 833 YTVPEATPDAGDDPKVTRAKFIRDLFLRISTATGDKHYCYPHTCVADTENNIRRVFN 892

QY 442 DCRDIIQRMHLKQYELL 458

DB 893 DCRDIIQRMHLKQYELL 909

RESULT 15

ABM82265

ID ABM82265 standard; protein; 909 AA.

AC ABM82265;

XX 18-NOV-2004 (first entry)

DE Tumour-associated antigenic target (TAT) polypeptide PRO83186, SEQ:5829.

XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;

XX tumour; diagnosis; cell proliferative disorder; breast cancer;

XX colorectal cancer; lung cancer; ovarian cancer; liver cancer;

XX central nervous system cancer; bladder cancer; pancreatic cancer;

XX cervical cancer; melanoma; leukaemia; hybridisation probe;

XX chromosome identification; chromosome mapping; gene mapping;

XX gene therapy; cytostatic.

XX Homo sapiens.

XX WO2004030615-A2.

XX 15-APR-2004.

XX 29-SEP-2003; 2003WO-US028547.

XX 02-OCT-2002; 2002US-0414971P.

XX (GETH ) GENENTECH INC.

XX Wu TD, Zhang Z, Zhou Y;

XX WPI; 2004-347921/32.

XX N-PSDB; ACN40807.



xx New tumor-associated antigenic target polypeptides and nucleic acids,  
 PT useful in preparing a medicament for treating or detecting a  
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
 PT prostate cancer or tumor.  
 xx Claim 12; SEQ ID NO 5829; 7273pp; English.  
 xx  
 CC The invention relates to human tumour-associated antigenic target (TAT)  
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
 CC overexpressed in cancer tissues compared to normal tissues, and may thus  
 CC serve as effective targets for the diagnosis and treatment of cancer in  
 CC mammals. The invention also relates to nucleic acid and polypeptide  
 CC sequences at least 80% identical to the TAT nucleic acids and  
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
 CC TAT polypeptide; and methods and compositions for the treatment or  
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
 CC antibodies, antagonists, binding molecules and compositions are useful  
 CC for diagnosing or treating a cell proliferative disorder associated with  
 CC increased TAT expression, particularly cancers such as breast cancer,  
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
 CC used as hybridisation probes, in chromosome and gene mapping, in  
 CC chromosome identification and in gene therapy. The present sequence  
 CC represents a TAT polypeptide of the invention  
 xx  
 SQ Sequence 909 AA;

Query Match 65.0%; Score 1559; DB 8; Length 909;  
 Best Local Similarity 68.6%; Pred. No. 2.9e-126;  
 Matches 300; Conservative 49; Mismatches 68; Indels 20; Gaps 4;  
 QY 42 RAARDTARTLLPR--GGESESPACAPKADKP-KEKQRTQEQLSAEEREAEREAVKE 97  
 DB 473 RKPQNLRLNFLVQAFGFCGRSESPQKASRSILKVKVPLAEKRQMRKEALEKRAQR 532  
 QY 98 A-RKVSIGDIRMLDDQRDLQOTHRLLLLGAGESGKSTIVKQMRILHVNQFNPE----- 150  
 DB 533 AEKRSKLIDKQLQDERXGMYCTHRLLLGAGESGKSTIVKQMRILHVNQFNQGEEDP 592  
 QY 151 -----EKKOKILDIRKNVDAIVTIVSAMSTIIPVPLANPENQFESDYIKSIAP 201  
 DB 593 QAARNSDGEKATKQDIKNLKEALITIVAASNLVPPVELANPENQFVDYILSMNV 652  
 QY 202 TDFEYSQEFEDHVKKLWDEGVKACFERSNEYQLIDCAQYFLERIDSVLSVDTPTDQDL 261  
 DB 653 PDFDFPPEFVEHAKALWEDEGVACVRSNEYQLIDCAQYFLDKIDVIKQADYVPSDQDL 712  
 QY 262 LRCRVLTSGIFETRFQVDKYNFMFDVGGQDERRKWIQCFNDVTATIIYVAACSSYNMVI 321  
 DB 713 LRCRVLTSGIFETRFQVDKYNFMFDVGGQDERRKWIQCFNDVTATIIYVASSSYNMVI 772  
 QY 322 REDNNTNRLRESLDLRFESIWNNRWLRTISILFNKODMLAEKVLACKSKIEDYFPEYAN 381  
 DB 773 REDNNTNRLQELNLFKSIWNNRWLRTISVILFNKODMLAEKVLACKSKIEDYFPEFAR 832  
 QY 382 YTVPEATPDAGDPKVKTRAKFFIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFN 441  
 DB 833 YTVPEATPDAGDPKVKTRAKFFIRDEFIRISTASGDGRHYCYPHFTCAVDTENIRRVFN 892  
 QY 442 DCRDIITQRMHLKQVELL 458  
 DB 893 DCRDIITQRMHLRQVELL 909

Search completed: March 2, 2006, 19:32:17  
 Job time : 145.364 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:30:27 ; Search time 27.1616 Seconds  
(without alignments)  
1622.414 Million cell updates/sec

Title: US-10-618-320A-1  
Perfect score: 2400  
Sequence: 1 MGLCYSLRLLFGPGDDPC.....VFNDCRDIQRMHLKQYELL 458  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed; and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1819	75.8	381	2 I53271	G-protein alpha-ol
2	1548	64.5	846	2 S52418	GTP-binding regula
3	1530	63.7	377	1 RGMSA1	GTP-binding regula
4	1529.5	63.7	380	1 RGHUA1	GTP-binding regula
5	1527	63.6	379	1 RGXLA	GTP-binding regula
6	1525.5	63.6	394	2 S33458	GTP-binding regula
7	1525	63.5	395	1 RGHUA2	GTP-binding regula
8	1522.5	63.4	394	1 RGMSA2	GTP-binding regula
9	1520.5	63.4	394	1 RGHVAE	GTP-binding regula
10	1520.5	63.4	394	1 RGHVA2	GTP-binding regula
11	1520.5	63.4	394	1 RGRTA2	GTP-binding regula
12	1518	63.2	397	1 RGPQA2	GTP-binding regula
13	1513.5	63.1	394	1 RBOGA	GTP-binding regula
14	1460	60.8	419	2 S34421	GTP-binding regula
15	1413	58.9	376	2 S27015	GTP-binding regula
16	1398	58.2	382	1 RGFPA5	GTP-binding regula
17	1397.5	58.2	335	2 JH0813	GTP-binding regula
18	1384.5	57.7	385	1 RGFPA1	GTP-binding regula
19	1292	53.8	375	2 T37245	GTP-binding regula
20	1278.5	53.3	385	2 D87723	protein R06A10.2
21	1270.5	52.9	379	2 A42964	guanine nucleotide
22	739.5	30.8	353	2 S34347	GTP-binding regula
23	725.5	30.2	355	2 T15288	GTP-binding regula
24	721	30.0	354	2 S27014	hypothetical prote
25	720.5	30.0	359	2 S45699	GTP-binding regula
26	719.5	30.0	359	1 RGM5Q	GTP-binding regula
27	717.5	29.9	359	2 S71963	GTP-binding protei
28	715	29.8	354	1 RGFPO1	GTP-binding regula
29	715	29.8	354	2 S24352	gustducin - rat

ALIGNMENTS

RESULT 1

I53271  
G-protein alpha-olf subunit - human  
C:Species: Homo sapiens (man)  
C:Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004  
C:Accession: I53271  
R:Zigman, J.M.; Westermarck, G.T.; LaMendola, J.; Boel, E.; Steiner, D.F.  
Endocrinology 133, 2508-2514, 1993  
A:Title: Human G(olf) alpha: complementary deoxyribonucleic acid structure and expression  
A:Reference number: I53271; MUID:94062616; PMID:8443272  
A:Accession: I53271  
A:Status: preliminary;  
A:Molecule type: mRNA  
A:Residues: 1-381 <RES>  
A:Cross-references: UNIPROT:P38405; UNIPARC:UPI0000001246; GB:L10665; NID:g182967; PIDN:;  
A:Superfamily: GTP-binding regulatory protein Gs alpha chain  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F:49-56/Region: nucleotide-binding motif A (P-loop)  
F:279-282/Region: GTP-binding NKXD motif

Query Match 75.8%; Score 1819; DB 2; Length 381;  
Best Local Similarity 95.0%; Pred. No. 1e-108;  
Matches 345; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY	96	KEARKVSRGIDRMRLRDKRDLOQTHRLLLLGAGESGKSTIVKQMRILHVNGFNPEKKOK	155
DB	19	KERREANKKIEKQLQKERLAYKATHRLLLLGAGESGKSTIVKQMRILHVNGFNPEKKOK	78
QY	156	ILDIRKNVKDAIVTVISANSTIIPVPLANPENQFRSDYIKSIAPITDPEYSQEFDFHVK	215
DB	79	ILDIRKNVKDAIVTVISANSTIIPVPLANPENQFRSDYIKSIAPITDPEYSQEFDFHVK	138
QY	216	KLWDEGVKACERSNEYQLIDCAQYFLERISVLSVDYTPDQDQLLRCLVLTSGIFETR	275
DB	139	KLWDEGVKACERSNEYQLIDCAQYFLERISVLSVDYTPDQDQLLRCLVLTSGIFETR	198
QY	276	FQVDKVNFMFMDVGGQDERRKWIQCFNDVTALIVVAACSSYNMVIREDNNTNRLRESLD	335
DB	199	FQVDKVNFMFMDVGGQDERRKWIQCFNDVTALIVVAACSSYNMVIREDNNTNRLRESLD	258
QY	336	LFESIWNNRWLRTISILFLNFKQDMLAEKVLGAKSKIEDYFPPEYANTVTPEDATPDAGD	395
DB	259	LFESIWNNRWLRTISILFLNFKQDMLAEKVLGAKSKIEDYFPPEYANTVTPEDATPDAGD	318
QY	396	PKVTRAKFFIRDLFRLRISTATGDKHKYCPHFTCAVDENIRRVFNDCRDIIQRMHLKQY	455
DB	319	PKVTRAKFFIRDLFRLRISTATGDKHKYCPHFTCAVDENIRRVFNDCRDIIQRMHLKQY	378
QY	456	ELL 458	
DB	379	ELL 381	

GTP-binding regula  
GTP-binding regula  
GTP-binding regula  
GTP-binding regula  
GTP-binding regula  
GTP-binding protei  
GTP-binding regula  
GTP-binding regula  
GTP-binding protei  
GTP-binding protei  
GTP-binding regula  
GTP-binding regula  
GTP-binding regula  
GTP-binding regula

## RESULT 2

S52418  
GTP-binding regulatory protein Gs alpha-XL chain - rat  
N;Alternate names: G protein XL-alpha-s  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 14-Jul-1995 #sequence\_revision 10-Nov-1995 #text\_change 02-Feb-2001  
C;Accession: S52418  
R;Kehlenbach, R.H.; Matthey, J.; Huttnner, W.B.  
Nature 372, 804-809, 1994  
A;Title: XL-alpha-s is a new type of G protein.  
F;132-846/Region: GTP-binding NKXD motif  
A;Reference number: S52418; MUID:95089824; PMID:7997272  
A;Accession: S52418  
A;Molecule type: mRNA  
A;Residues: 1-846 <KEH>  
A;Cross-references: UNIPARC:UPI000017C91F; EMBL:X84047; NID:G642267; PIDN:CAA58866.1; PI

R;Kehlenbach, R.H.; Matthey, J.; Huttnner, W.B.  
Nature 375, 253, 1995  
A;Title: Correction: XLalphas is a new type of G protein.  
A;Reference number: S58911  
A;Contents: annotation; assignment of start\_codon  
A;Note: experimental data from this paper suggest that the translation is initiated at P

F;132-846/Product: GTP-binding regulatory protein Gs alpha-XL chain #status experimental  
F;499-506/Region: nucleotide-binding motif A (P-loop)  
F;744-747/Region: GTP-binding NKXD motif

Query Match 64.5%; Score 1548; DB 2; Length 846;  
Best Local Similarity 70.2%; Pred. No. 5.3e-91;  
Matches 292; Conservative 47; Mismatches 61; Indels 16; Gaps 2;

Qy	59	GSFACAPKADKP-KEKQRTQEQLSAEEREAAREAVKARVSGIDRMRLDQKRDQ 117
Db	431	GLSECTRSLSGKADKPMEERKQMRKEAMERQKADKRSKLIDKQLEEEKQDYM 490
Qy	118	QTHRLLLGAGESGKSTIVKQMRILHVNGFNPE-----EKKQKILDIRKN 162
Db	491	CTHRLLLGAGESGKSTIVKQMRILHVNGFNGEGBEDPQAARSNDGEKATKQVQIKNN 550
Qy	163	VKDAIVTVSAMSTIIPVPLANPENQFRSDYTKSAPITDFPSYQEFDPDHVKLLWDEG 222
Db	551	LKEAETIVAAMSLVPPVELANPENQFRVDYILSVNMVNFDFPFPEFVEHAKALWEDEG 610
Qy	223	VKACFERSNEYQLDCAQYFLERIDSVLVDYPTDQDLLRCRLVLTSGIFETRFQVDKN 282
Db	611	VRACYERSNEYQLDCAQYFLDKIDVIKQADYVPSDQDLPRCRLVLTSGIFETKPFQVDKN 670
Qy	283	FHMFVGGQDERRKWTQCFNDVTALIIYVAACSSYNMVIREDNNTNRLRESLDLPESIWN 342
Db	671	FHMFVGGQDERRKWTQCFNDVTALIIYVASSSYNMVIREDNQTNRLQEQALNLFKSIWN 730
Qy	343	NRWLRTISILFLNKQDMLAEKVLGSKIEDYFPEYANYTVPEPATPDAGEDPKVTRAK 402
Db	731	NRWLRTISVLFLNKQDMLAEKVLGSKIEDYFPEFARYTTPEDATPEPGEDPRVTRAK 790
Qy	403	FFIRDLPLRISTATGDKGHYCPHFCAVDTEINRRVFNDCRDIIOHMLKQYELL 458
Db	791	YFIRDFLRISTASGDRGHYCPHFCAVDTEINRRVFNDCRDIIOHMLRQYELL 846

## RESULT 3

RGMSAL  
GTP-binding regulatory protein Gs alpha-S1 chain (adenylate cyclase-stimulating) - mouse  
N;Alternate names: guanine nucleotide binding protein Gs alpha-S1 chain; heterotrimeric  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
C;Accession: A25889

R;Sullivan, K.A.; Liao, Y.C.; Alborzi, A.; Beiderman, B.; Chang, F.H.; Masters, S.B.; Le  
Proc. Natl. Acad. Sci. U.S.A. 83, 6687-6691, 1986  
A;Title: Inhibitory and stimulatory G proteins of adenylate cyclase: cDNA and amino acid  
A;Reference number: A94123; MUID:86313643; PMID:3092218  
A;Accession: A25889  
A;Molecule type: mRNA

A;Residues: 1-377 <SUL>

A;Cross-references: UNIPROT:P04894; UNIPARC:UPI000016CDAC; GB:M13964; NID:G193652; PIDN:Y  
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay  
ains. The beta and gamma chains, required for GTPase activity, appear to be common to all  
ase; it is specific for each type of G protein.

C;Comment: The Gs alpha chain is specific for G protein that is responsible for transduct  
xigic stimuli.

C;Superfamily: GTP-binding regulatory protein Gs alpha chain

C;Keywords: alternative splicing; GTP binding; heterotrimer; nucleotide binding; P-loop;  
F;45-52/Region: nucleotide-binding motif A (P-loop)

F;275-278/Region: GTP-binding NKXD motif

F;51/Binding site: GTP (Lys) #status predicted

F;184/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 63.7%; Score 1530; DB 1; Length 377;

Best Local Similarity 76.9%; Pred. No. 2.7e-90;

Matches 279; Conservative 45; Mismatches 39; Indels 0; Gaps 0;

Qy	96	KEARKVSRGIDRMRLDQKRDLOQTHRLLLLGAGESGKSTIVKQMRILHVNGFNPEKKOK 155
Db	15	KAQREANKKIEKQKQDKQVYRATHRLLLLGAGESGKSTIVKQMRILHVNGFNGBDEKATK 74

Qy	156	ILDIRKNVKDAIVTVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQEPFHVK 215
Db	75	VQDIKNLKEAETIVAAMSLVPPVELANPENQFRVDYILSVMLPNDFPFPEFVEHAK 134

Qy	216	KLWDEGVKACFERSNEYQLDCAQYFLERIDSVSLVDYPTDQDLLRCRLVLTSGIFETR 275
Db	135	ALWDEGVKACFERSNEYQLDCAQYFLDKIDVIKQADYVPSDQDLLRCRLVLTSGIFETR 194

Qy	276	FOVDKVNFMFVGGQDERRKWTQCFNDVTALIIYVAACSSYNMVIREDNNTNRLRESLD 335
Db	195	FQVDKVNFMFVGGQDERRKWTQCFNDVTALIIYVASSSYNMVIREDNQTNRLQEQALN 254

Qy	336	LFESINNRWLRTISILFLNKQDMLAEKVLGSKIEDYFPEYANYTVPEPATPDAGED 395
Db	255	LFKSIWNRWLRTISVLFLNKQDMLAEKVLGSKIEDYFPEFARYTTPEDATPEPGED 314

Qy	396	PKVTRAKFIRDLPLRISTATGDKGHYCPHFCAVDTEINRRVFNDCRDIIOHMLKQY 455
Db	315	PRVTRAKFIRDEFRLRISTASGDRGHYCPHFCAVDTEINRRVFNDCRDIIOHMLRQY 374

Qy	456	ELL 458
Db	375	ELL 377

## RESULT 4

RGHUAL

GTP-binding regulatory protein Gs alpha chain (adenylate cyclase-stimulating), splice for  
N;Alternate names: guanine nucleotide binding protein Gs alpha-S1 chain; heterotrimeric  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 31-Dec-2004  
C;Accession: C31927; D31927; B24366

R;Kozasa, T.; Itoh, H.; Tsukamoto, T.; Kaziro, Y.

Proc. Natl. Acad. Sci. U.S.A. 85, 2081-2085, 1988

A;Title: Isolation and characterization of the human G-s-alpha gene.

A;Reference number: A31927; MUID:88176890; PMID:3127824

A;Accession: C31927

A;Molecule type: DNA

A;Residues: 1-71,73-380 <KQZ3>

A;Cross-references: UNIPROT:P04895; UNIPROT:Q14433; UNIPARC:UPI000004D203; GB:M21142; DDI

A;Note: splice form 3

A;Accession: D31927

A;Molecule type: DNA

A;Residues: 1-380 <KQZ4>

A;Cross-references: UNIPARC:UPI00000001248; GB:M21142; DDBJ:J03647; NID:G183402; PIDN:AAA

A;Note: splice form 4

R;Mattera, R.; Codina, J.; Crozat, A.; Kidd, V.; Woo, S.L.C.; Birnbaumer, L.

FEBS Lett. 206, 36-42, 1986

A;Title: Identification by molecular cloning of two forms of the alpha-subunit of the hu

A;Reference number: A24366; MUID:87005246; PMID:3093273

```

rase; it is specific for each type of G protein.
C:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprot
F:2-379/Product: GTP-binding regulatory protein Gs alpha chain #status predicted <MAT>
F:47-54/Region: nucleotide-binding motif A (P-loop)
F:277-280/Region: GTP-binding NKXD motif
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F:3/Binding site: palmitate (Cys) (covalent) #status predicted
F:186/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match          63.6%; Score 1527; DB 1; Length 379;
Best Local Similarity 76.4%; Pred. No. 4.2e-90;
Matches 281; Conservative 43; Mismatches 44; Indels 0; Gaps 0;

QY      91 EREAIVEAKVSRGIDRMLRDQKRDLOOQTHRLLLLGAGESGKSTIVKQMRILHVNFGNPE 150
       :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      12 QRNEEKVQRETNNKKIEQLQDKQKVRYATHRLLLLGAGESGKSSIVKQMRILHVNFGNAE 71

QY     151 EKKOKILDIRKNVKDAITTVSAMSITIIPVPLANENQFRSDYIKSIAPITDFEYSQEF 210
       :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      72 EKTKVQDQIKNNIKEAETITVTAGNLSPPVVELNPENQFRIDYILNLPNYKDFEPSPEF 131

QY     211 FDHVKLWDDEGVKACERNSEYQLIDCAQYFLERIDSLSVDYPTTDODLLRCRVLTSG 270
       :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB     132 YEHTKLWQDEGRACYERSNEYQLIDCAQYFLKDIDVKQNDYTPSDQDLRCRVLTSG 191

QY     271 IPETRFQVDKVFPHFDVGQORDERRRWIQCFNDVTAIYYVAACSSYNMVIREDNNTNRL 330
       :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB     192 IFETKFQVDKVFPHFDVGQORDERRRWIQCFNDVTALIFVVASSYNMVIREDNNTNRL 251

QY     331 RESLDLFSIWNRRWLRTISILFLNKQDMLAELKVLGACKSKIEDYFPPEYANTVPEDATP 390
       :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB     252 QBALNFLFKSIWNRRWLRTISVILFLNKQDLLEAKVGNACKSKIEDYFPPEARVTTTPDATP 311

QY     391 DAGEDPKVTRAKFIIRDLFRISTATGDGKYCYCPHFCAVDTENIRRVFNDCRDDIIORM 450
       :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB     312 EVGEDPRVTRAKFYIRDEFIRISTASGGHRHCYPHFCAVDTENIRRVFNDCRDDIIORM 371

QY     451 HLKQYELL 458
       :|:::|:::|
DB     372 HLKQYELL 379

RESULT 6
S33458
GTP-binding regulatory protein Gs alpha chain - dog
C:Species: Canis lupus familiaris (dog)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S33458
R:Ishikawa, Y.; Homcy, C.J.
submitted to the EMBL Data Library, June 1992
A:Description: cDNA sequence for alpha subunit of stimulatory guanine nucleotide binding
A:Reference number: S33458
A:Accession: S33458
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-394 <ISH>
A:Cross-references: UNIPROT:P04895; UNIPARC:UPI0000001247; EMBL:Z12168; NID:g311336; PID:
C:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: GTP binding; nucleotide binding; P-loop
F:47-54/Region: nucleotide-binding motif A (P-loop)
F:292-295/Region: GTP-binding NKXD motif

Query Match          63.6%; Score 1525.5; DB 2; Length 394;
Best Local Similarity 73.6%; Pred. No. 5.4e-90;
Matches 282; Conservative 44; Mismatches 42; Indels 15; Gaps 1;

QY      91 EREAIVEAKVSRGIDRMLRDQKRDLOOQTHRLLLLGAGESGKSTIVKQMRILHVNFGNPE 150
       :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB      12 QRNEEKAQREANKKIEQLQDKQKVRYATHRLLLLGAGESGKSTIVKQMRILHVNFGNCE 71

QY     151 -----EKKQKILDIRKNVKDAITTVSAMSTIIPVPLANPENQFRSDYI 195

```

Db 72 GGEDPQAARSNDGEKATKVQDIKNLKEAETIVAAMSNLVPVVELANPENQFRVDYI 131  
QY 196 KSTAPITDFEYSQEFFDHVKKLWDEGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYT 255  
Db 132 LSVNVPDFPFPEFYEHAKLWDEGVKACFERSNEYQLIDCAQYFLDKIDVIKQADYV 191  
QY 256 PTQDLRLCRVLTSIGIFETRFQVDKNVFNHFDVGGQDRERRKWIQCFNDVTALIIYVAACS 315  
Db 192 PSODDLRLCRVLTSIGIFETRFQVDKNVFNHFDVGGQDRERRKWIQCFNDVTALIIYV 251  
QY 316 SYNVMVREDNNTLRSLDLFESIMNRLRTISILFLNKKQDLAEKVLAKGSKIETY 375  
Db 252 SYNVMVREDNNTLRSLDLFESIMNRLRTISILFLNKKQDLAEKVLAKGSKIETY 311  
QY 376 FPEVANYTVPEADPDAGEDPKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTEN 435  
Db 312 FPEFARYTTPEDATPEGEDPRVTRAKYFIRDFLRISTATGDKGHYCYPHFTCAVDTEN 371  
QY 436 IRRVFNDCRDIIRQMHKLQYELL 458  
Db 372 IRRVFNDCRDIIRQMHKLQYELL 394

RESULT 7  
RGHUA2  
GTP-binding regulatory protein Gs alpha chain (adenylate cyclase-stimulating), splice fo  
N;Alternate names: guanine nucleotide binding protein Gs alpha-S2 chain; heterotrimeric  
N;Contains: GTP-binding regulatory protein Gs alpha chain, splice form 1  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1990 #sequence\_revision 15-Jun-1996 #text\_change 09-Jul-2004  
C;Accession: B31927; A24366; S02122; A25919  
R;Kozasa, T.; Itoh, H.; Teukamoto, T.; Kaziro, Y.  
Proc. Natl. Acad. Sci. U.S.A. 85, 2081-2085, 1988  
A;Title: Isolation and characterization of the human G-s-alpha gene.  
A;Reference number: A31927; MUID:98176890; PMID:3127824  
A;Accession: B31927  
A;Molecule type: DNA  
A;Residues: 1-395 <KOZ2>  
A;Cross-references: UNIPROT:O96H70; UNIPARC:UPI0000004D202; GB:M21142; DDBJ:J03647; NID:9  
A;Note: splice form 2  
A;Accession: A31927  
A;Molecule type: DNA  
A;Residues: 1-86,88-395 <KOZ1>  
A;Cross-references: UNIPARC:UPI0000001247; GB:M21142; DDBJ:J03647; NID:g183402; PIDN:AAA  
A;Note: splice form 1  
R;Mattera, R.; Codina, J.; Crosat, A.; Kidd, V.; Woo, S.L.C.; Birnbaumer, L.  
FEBS Lett. 206, 36-42, 1986  
A;Title: Identification by molecular cloning of two forms of the alpha-subunit of the hu  
A;Reference number: A24366; MUID:87005246; PMID:3093273  
A;Accession: A24366  
A;Molecule type: mRNA  
A;Residues: 1-86,88-395 <MAT>  
A;Cross-references: UNIPARC:UPI0000001247; EMBL:X04408; NID:g31914; PIDN:CAA27996.1; PID  
R;Harris, B.A.  
Nucleic Acids Res. 16, 3585, 1988  
A;Title: Complete cDNA sequence of a human stimulatory GTP-binding protein alpha subunit  
A;Reference number: S02122; MUID:88233954; PMID:3131741  
A;Accession: S02122  
A;Molecule type: mRNA  
A;Residues: 1-5,'T',7-86,88-395 <HAR>  
A;Cross-references: UNIPARC:UPI000016AA01; EMBL:X07036; NID:g31951; PIDN:CAA30084.1; PID  
R;Bray, P.; Carter, A.; Simons, C.; Guo, V.; Puckett, C.; Kamholz, J.; Spiegel, A.; Nire  
Proc. Natl. Acad. Sci. U.S.A. 83, 8893-8897, 1986  
A;Title: Human cDNA clones for two species of G-alpha-s-signal transduction protein.  
A;Reference number: A25919; MUID:87067419; PMID:3024154  
A;Accession: A25919  
A;Molecule type: mRNA  
A;Residues: 12-230,'Q',232-395 <BRA>  
A;Cross-references: UNIPARC:UPI000016A9C8; GB:M14631; NID:g183416; PIDN:AAAS2583.1; PID:  
A;Note: The authors translated the codon CAA for residue 221 as Glu  
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay  
ains. The beta and gamma chains, required for GTPase activity, appear to be common to al  
rase; it is specific for each type of G protein.

C;Comment: The Gs alpha chain is specific for G protein that is responsible for transduct  
rgic stimuli.  
C;Comment: See also PIR:RGHUA1.  
C;Genetics:  
A;Gene: GDB:GNAS1; GNAS  
A;Cross-references: GDB:120628; OMIM:139320  
A;Map position: 20q13.2-20q13.3  
A;Introns: 47/1; 71/2; 86/2; 105/3; 145/3; 178/2; 196/3; 221/2; 241/1; 281/2; 325/1; 347/  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
C;Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprote  
F;2-395/Product: GTP-binding regulatory protein Gs alpha chain, splice form 2 #status pr  
F;2-86,88-395/Product: GTP-binding regulatory protein Gs alpha chain, splice form 1 #stat  
F;47-54/Region: nucleotide-binding motif A (P-loop)  
F;293-296/Region: GTP-binding NKXD motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;3/Binding site: myristate (Cys) (covalent) #status predicted  
F;202/Modified site: ADP-riboylarginine (Arg) (by cholera toxin) #status predicted

Query Match 63.5%; Score 1525; DB 1; Length 395;  
Best Local Similarity 73.4%; Pred. No. 5.9e-90;  
Matches 282; Conservative 44; Mismatches 42; Indels 16; Gaps 1;

QY 91 EREAVKEARKVSGIDRMRLRQKRDQLQOTHRLLLLGAGSGKSTIVKQMRILHVNQFNPE 150  
Db 12 QRNEEKAQREANKTEKQLQKQVYRATHRLLLLGAGSGKSTIVKQMRILHVNQFNGE 71  
QY 151 -----EKQKILDIRKNVKDAIVTVSAMSTLIIPVPLANENQFRSDY 194  
Db 72 GGEEDPQAARSNDGSEKATKVQDIKNLKEAETIVAAMSNLVPVVELANPENQFRVDY 131  
QY 195 IKSIAPIITDFEYSQEFFDHVKKLWDEGVKACFERSNEYQLIDCAQYFLERIDSVSLVDY 254  
Db 132 ILSVMNVPDFPFPEFYEHAKLWDEGVKACFERSNEYQLIDCAQYFLDKIDVIKQADY 191  
QY 255 TPTQDLRLCRVLTSIGIFETRFQVDKNVFNHFDVGGQDRERRKWIQCFNDVTALIIYVAAC 314  
Db 192 VPSQDLRLCRVLTSIGIFETRFQVDKNVFNHFDVGGQDRERRKWIQCFNDVTALIIFWAS 251  
QY 315 SYNVMVREDNNTLRSLDLFESIMNRLRTISILFLNKKQDLAEKVLAKGSKI 374  
Db 252 SYNVMVREDNNTLRSLDLFESIMNRLRTISILFLNKKQDLAEKVLAKGSKI 311  
QY 375 YPEFANYTVPEADPDAGEDPKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTE 434  
Db 312 YPEFARYTTPEDATPEGEDPRVTRAKYFIRDFLRISTATGDKGHYCYPHFTCAVDTE 371  
QY 435 NIRRVFNDCRDIIRQMHKLQYELL 458  
Db 372 NIRRVFNDCRDIIRQMHKLQYELL 395

RESULT 8  
RGM5A2  
GTP-binding regulatory protein Gs alpha-S2 chain (adenylate cyclase-stimulating) - mouse  
N;Alternate names: guanine nucleotide binding protein Gs alpha-S2 chain; heterotrimeric  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 19-Jan-2001  
C;Accession: S03075  
R;Rall, T.; Harris, B.A.  
FEBS Lett. 224, 365-371, 1987  
A;Title: Identification of the lesion in the stimulatory GTP-binding protein of the unco  
A;Reference number: S03075; MUID:88083563; PMID:2826231  
A;Accession: S03075  
A;Molecule type: mRNA  
A;Residues: 1-394 <RAL>  
A;Cross-references: UNIPARC:UPI0000054C7A; GB:Y00703; NID:g51127; PIDN:CAA68695.1; PID:g  
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay  
ains. The beta and gamma chains, required for GTPase activity, appear to be common to al  
rase; it is specific for each type of G protein.  
C;Comment: The Gs alpha chain is specific for G protein that is responsible for transduct  
rgic stimuli.  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
C;Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprote

Qy	91	EREAVTEARKVSRGIDRMRLRDQKDLQOTHRLLLLGAGESGKSTIVKQMRILHVNGFNPE	150
Db	12	ORNEEKAQREANKKIEKQLQDKQVYRATHRLLLLGAGESGKSTIVKQMRILHVNGFNGE	71
Qy	151	-----EKKQKILDIRKNVKDAIVTIVSAMSTIIPVPPLANPENQFRSDYI	195
Db	72	GGEEDPOAARNSDGEKATKVQDINKNLKEAETIIVAAMSNLVPVELANPENQFRVDYI	131
Qy	196	KSIAPITDFEYSQBFDFHVKKLWDEGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYT	255
Db	132	LSVMNVNFDPPFYEYHAKALWEDEGVACYSRNEYQLIDCAQYFLDKIDVKKQADV	191
Qy	256	PTDQDLRCRVLTSGIIPETRFQVDKVNPFMDVGQQRDERKRWIQCFNDVTAILIYVAACS	315
Db	192	PSDQDLRCRVLTSGIIFETKFQVDKVNPFMDVGQQRDERKRWIQCFNDVTAILIIFVWASS	251
Qy	316	SYNVIREDNNTNLRISLDFESIWNNRWLRTTSIIILFLNKKQDMLAEKVLGKSKIEDY	375
Db	252	SYNVIREDQTNRLQBALNLFPSIWNNRWLRTTSIIILFLNKKQDMLAEKVLGKSKIEDY	311
Qy	376	FPEYANTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDKGHYCVPHFTCAVDTEN	435
Db	312	FPEFARVTTPEDATPEGEDPRVTRAKYFIRDEFRLRISTATSGDGRHYCVPHFTCAVDTEN	371
Qy	436	IRRVFNDCRDIIORMHLKQYELL	458
Db	372	IRRVFNDCRDIIORMHLRQYELL	394
RESULT 10			
RGHYA2			
GTP-binding regulatory protein Gs alpha-2 chain (adenylate cyclase-stimulating			
N;Alternate names: guanine nucleotide binding protein Gs alpha-2 chain; heterotrimeric Gs; Mesocricetus auratus (golden hamster)			
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004			
C;Accession: S10508			
R;Conner, D.A.; Feldman, A.M.; van Dop, C.			
Nucleic Acids Res. 18, 4279, 1990			
A;Title: cDNA sequence for the alpha subunit of the guanine nucleotide-binding			
A;Reference number: S10508; MUID:90332451; PMID:2115997			
A;Accession: S10508			
A;Molecule type: mRNA			
A;Residues: 1-394 <CON>			
A;Cross-references: UNIPROT:P04894; UNIPARC:UPI0000001ACB; EMBL:X53139; NID:94545			
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that			
ains. The beta and gamma chains, required for GTPase activity, appear to be con-			
rase; it is specific for each type of G protein.			
C;Comment: The Gs alpha chain is specific for G protein that is responsible for			
ergic stimuli.			
C;Superfamily: GTP-binding regulatory protein Gs alpha chain			
C;Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer			
F;2-394/Product: GTP-binding regulatory protein Gs alpha-2 chain #status predicted			
F;47-54/Region: nucleotide-binding motif A (P-loop)			
F;292-295/Region: GTP-binding NKXD motif			
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted			
F;3/Binding site: palmitate (Cys) (covalent) #status predicted			
F;201/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted			
Query Match 63.4%; Score 1520.5; DB 1; Length 394;			
Best Local Similarity 73.4%; Pred. No. 1.1e-89;			
Matches 281; Conservative 45; Mismatches 42; Indels 15; Gaps 1			
Qy	91	EREAVTEARKVSRGIDRMRLRDQKDLQOTHRLLLLGAGESGKSTIVKQMRILHVNGFNPE	150
Db	12	ORNEEKAQREANKKIEKQLQDKQVYRATHRLLLLGAGESGKSTIVKQMRILHVNGFNGE	71
Qy	151	-----EKKQKILDIRKNVKDAIVTIVSAMSTIIPVPPLANPENQFRSDYI	195
Db	72	GGEEDPOAARNSDGEKATKVQDINKNLKEAETIIVAAMSNLVPVELANPENQFRVDYI	131
Qy	196	KSIAPITDFEYSQBFDFHVKKLWDEGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYT	255



QY 256 PTQDRLRCRVLTSGIFETKFOVDKVNPHMFVGGQDERRKWIQCFNDVTAIIYVAACS 315  
 Db 192 PSQDRLRCRVLTSGIFETKFOVDKVNPHMFVGGQDERRKWIQCFNDVTAIIYVASS 251  
 QY 316 SYNWVIREDDNNRLRESLDLFIWNNRRLRTISILFLNKQDMLAEKVLGKSKIEDY 375  
 Db 252 SYNWVIREDDNNRLRESLDLFIWNNRRLRTISILFLNKQDMLAEKVLGKSKIEDY 311  
 QY 376 PPEYANVTVPEDATPDAGDPKVTAKFPFIRDLPLRISTATGDKHKYCYPHFTCAVDTEN 435  
 Db 312 PPEYANVTVPEDATPDAGDPKVTAKFPFIRDLPLRISTATGDKHKYCYPHFTCAVDTEN 371  
 QY 436 IRRVFNDCRDIIOHMLKQYELL 458  
 Db 372 IRRVFNDCRDIIOHMLKQYELL 394

RESULT 11  
 RGR7A2  
 GTP-binding regulatory protein Gs alpha-2 chain (adenylate cyclase-stimulating) - rat  
 N;Alternate names: guanine nucleotide binding protein Gs alpha-2 chain; heterotrimeric G  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
 C;Accession: A27423; C24882  
 R;Jones, D.T.; Reed, R.R.  
 J. Biol. Chem. 262, 14241-14249, 1987  
 A;Title: Molecular cloning of five GTP-binding protein cDNA species from rat olfactory r  
 A;Reference number: A92614; MUID:88007678; PMID:2820999  
 A;Accession: A27423  
 A;Molecule type: mRNA  
 A;Residues: 1-394 <ON>  
 A;Cross-references: UNIPROT:P04894; UNIPARC:UPI0000001ACB; GB:M17525; NID:g203171; PIDN:  
 R;Itoh, H.; Kozasa, T.; Nagata, S.; Nakamura, S.; Katada, T.; Ui, M.; Iwai, S.; Ohseuka,  
 Proc. Natl. Acad. Sci. U.S.A. 83, 3776-3780, 1986  
 A;Title: Molecular cloning and sequence determination of cDNAs for alpha subunits of the  
 A;Reference number: A94707; MUID:86233317; PMID:3086867  
 A;Accession: C24882  
 A;Molecule type: mRNA  
 A;Residues: 1-394 <TO>  
 A;Cross-references: UNIPARC:UPI0000001ACB; GB:M12673; NID:g204441; PIDN:AAA1261.1; PID:  
 C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay  
 ains. The beta and gamma chains, required for GTPase activity, appear to be common to al  
 rase; it is specific for each type of G protein.  
 C;Comment: The Gs alpha chain is specific for G protein that is responsible for transduc  
 rgic stimuli.  
 C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
 C;Keywords: blocked amino end; GTP binding; lipoprotein; myristylation; nucleotide bindi  
 F;2-394/Product: GTP-binding regulatory protein Gs alpha-2 chain #status predicted <MAT>  
 F;47-54/Region: nucleotide-binding motif A (P-loop)  
 F;232-295/Region: GTP-binding NKXD motif  
 F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F;3/Binding site: palmitate (Cys) (covalent) #status predicted  
 F;201/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 63.4%; Score 1520.5; DB 1; Length 394;  
 Best Local Similarity 73.4%; Pred. No. 1.1e-89;  
 Matches 281; Conservative 45; Mismatches 42; Indels 15; Gaps 1;

QY 91 EREAVKEARKVSRGIDRMRLDQKRDLOQTHRLLLLGAGESGKSTIVKQMRILHVGNGNPE 150  
 Db 12 QRNEEKAQREANKKIEKQKQVYRATHRLLLLGAGESGKSTIVKQMRILHVGNGN 71  
 QY 151 EKKQKILDIRKVKDAIVTVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQEF 210  
 Db 72 EKATKQDQIKNNLKEAETIVAAMSLNPPVELANPENQFRVDYILSVNVPDPPPEF 131  
 QY 211 FPHVKLWDEGVKACFERSNEYQLIDCAQYFLERIDSVLVDYTFDQDRLRCRVLTSG 270  
 Db 132 YEHALWDEGVKACFERSNEYQLIDCAQYFLERIDSVLVDYTFDQDRLRCRVLTSG 191  
 QY 271 IFETRFQVDKVNPHMFVGGQDERRKWIQCFNDVTAIIYVAACSNNVIREDDNNRL 330  
 Db 192 IFETRFQVDKVNPHMFVGGQDERRKWIQCFNDVTAIIYVASSNNVIREDDNNRL 251  
 QY 331 RESLDLFIWNNRRLRTISILFLNKQDMLAEKVLGKSKIEDY 372  
 Db 252 QEALNLFKSIWNNRRLRTISILFLNKQDMLAEKVLGKSKIELFVLDRLRFQERPF 311  
 QY 373 EDYFPYANVTVPEDATPDAGDPKVTAKFPFIRDLPLRISTATGDKHKYCYPHFTCAVD 432  
 Db 312 EDYFPYANVTVPEDATPDAGDPKVTAKFPFIRDLPLRISTATGDKHKYCYPHFTCAVD 371  
 QY 433 TENIRRVFNDCRDIIOHMLKQYELL 458

Db 192 PSQDRLRCRVLTSGIFETKFOVDKVNPHMFVGGQDERRKWIQCFNDVTAIIYVASS 251  
 QY 316 SYNWVIREDDNNRLRESLDLFIWNNRRLRTISILFLNKQDMLAEKVLGKSKIEDY 375  
 Db 252 SYNWVIREDDNNRLRESLDLFIWNNRRLRTISILFLNKQDMLAEKVLGKSKIEDY 311  
 QY 376 PPEYANVTVPEDATPDAGDPKVTAKFPFIRDLPLRISTATGDKHKYCYPHFTCAVDTEN 435  
 Db 312 PPEYANVTVPEDATPDAGDPKVTAKFPFIRDLPLRISTATGDKHKYCYPHFTCAVDTEN 371  
 QY 436 IRRVFNDCRDIIOHMLKQYELL 458  
 Db 372 IRRVFNDCRDIIOHMLKQYELL 394

RESULT 12  
 RGFCA2  
 GTP-binding regulatory protein Gs alpha-2 chain (adenylate cyclase-stimulating) - pig  
 N;Alternate names: guanine nucleotide binding protein Gs alpha-2 chain; heterotrimeric G  
 C;Species: Sus scrofa domestica (domestic pig)  
 C;Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
 C;Accession: S18963  
 R;Roth, D.A.; Kay, R.A.M.; Hammond, H.K.  
 submitted to the EMBL Data Library, January 1992  
 A;Reference number: S18963  
 A;Accession: S18963  
 A;Molecule type: mRNA  
 A;Residues: 1-397 <ROT>  
 A;Cross-references: UNIPROT:P29797; UNIPARC:UPI000012B208; EMBL:X63893; NID:g1957; PIDN:(  
 C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay  
 ains. The beta and gamma chains, required for GTPase activity, appear to be common to al  
 rase; it is specific for each type of G protein.  
 C;Comment: The Gs alpha chain is specific for G protein that is responsible for transduc  
 rgic stimuli.  
 C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
 C;Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprote  
 F;2-397/Product: GTP-binding regulatory protein Gs alpha-2 chain #status predicted <MAT>  
 F;47-54/Region: nucleotide-binding motif A (P-loop)  
 F;277-280/Region: GTP-binding NKXD motif  
 F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F;3/Binding site: palmitate (Cys) (covalent) #status predicted  
 F;186/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 63.2%; Score 1518; DB 1; Length 397;  
 Best Local Similarity 72.8%; Pred. No. 1.7e-89;  
 Matches 281; Conservative 45; Mismatches 42; Indels 18; Gaps 1;

QY 91 EREAVKEARKVSRGIDRMRLDQKRDLOQTHRLLLLGAGESGKSTIVKQMRILHVGNGNPE 150  
 Db 12 QRNEEKAQREANKKIEKQKQVYRATHRLLLLGAGESGKSTIVKQMRILHVGNGN 71  
 QY 151 EKKQKILDIRKVKDAIVTVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQEF 210  
 Db 72 EKATKQDQIKNNLKEAETIVAAMSLNPPVELANPENQFRVDYILSVNVPDPPPEF 131  
 QY 211 FPHVKLWDEGVKACFERSNEYQLIDCAQYFLERIDSVLVDYTFDQDRLRCRVLTSG 270  
 Db 132 YEHALWDEGVKACFERSNEYQLIDCAQYFLERIDSVLVDYTFDQDRLRCRVLTSG 191  
 QY 271 IFETRFQVDKVNPHMFVGGQDERRKWIQCFNDVTAIIYVAACSNNVIREDDNNRL 330  
 Db 192 IFETRFQVDKVNPHMFVGGQDERRKWIQCFNDVTAIIYVASSNNVIREDDNNRL 251  
 QY 331 RESLDLFIWNNRRLRTISILFLNKQDMLAEKVLGKSKIEDY 372  
 Db 252 QEALNLFKSIWNNRRLRTISILFLNKQDMLAEKVLGKSKIELFVLDRLRFQERPF 311  
 QY 373 EDYFPYANVTVPEDATPDAGDPKVTAKFPFIRDLPLRISTATGDKHKYCYPHFTCAVD 432  
 Db 312 EDYFPYANVTVPEDATPDAGDPKVTAKFPFIRDLPLRISTATGDKHKYCYPHFTCAVD 371  
 QY 433 TENIRRVFNDCRDIIOHMLKQYELL 458



Db 372 TENIRRVFNDRCRIIORMHLRQYELL 397

## RESULT 13

## RBOGA

GTP-binding regulatory protein Gs alpha-2 chain (adenylate cyclase-stimulating) - bovine  
N;Alternate names: guanine nucleotide binding protein Gs alpha-2 chain; heterotrimeric G  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 09-Jul-2004  
C;Accession: A23818; A23615; A23813  
R;Robshaw, J.D.; Russell, D.W.; Harris, B.A.; Smigel, M.D.; Gilman, A.G.  
Proc. Natl. Acad. Sci. U.S.A. 83, 1251-1255, 1986  
A;Title: Deduced primary structure of the alpha subunit of the GTP-binding stimulatory P  
A;Reference number: A23813; MUID:86149283; PMID:3081893  
A;Accession: A23818  
A;Molecule type: mRNA  
A;Residues: 1-394 <RBO>  
C;Cross-references: UNIPROT:P04896; UNIPARC:UPI000016C31A; GB:M13006; NID:g161127; PIDN:  
R;Nukada, T.; Tanabe, T.; Takahashi, H.; Noda, M.; Hirose, T.; Inayama, S.; Numa, S.  
FEBS Lett. 195, 220-224, 1986  
A;Title: Primary structure of the alpha-subunit of bovine adenylate cyclase-stimulating  
A;Reference number: A23615; MUID:86108870; PMID:3080331  
A;Accession: A23615  
A;Molecule type: mRNA  
A;Residues: 1-17, 'A', 19-84, 'D', 86-362, 'F', 364-394 <NUK>  
C;Cross-references: UNIPARC:UPI000012B204; GB:M13006  
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay  
a signal. The beta and gamma chains, required for GTPase activity, appear to be common to all  
types; it is specific for each type of G protein.  
C;Comment: The Gs alpha chain is specific for G protein that is responsible for transduc  
tion stimuli.  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
C;Keywords: alternative splicing; blocked amino end; GTP binding; heterotrimer; lipoprot  
F;2-394/Product: GTP-binding regulatory protein Gs alpha-2 chain #status predicted <MAT>  
F;47-54/Region: nucleotide-binding motif A (P-loop)  
F;292-295/Region: GTP-binding NKXD motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;3/Binding site: palmitate (Cys) (covalent) #status predicted  
F;201/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 63.18; Score 1513.5; DB 1; Length 394;  
Best Local Similarity 73.48; Pred. No. 3.2e-89; Indels 15; Gaps 1;  
Matches 281; Conservative 44; Mismatches 43;  
Qy 91 EBAVKEARKVSRGIDRMLDQKRDLOOQTHRLLLGAGESGKSTIVKQMLHVNFGNPE 150  
Db 12 QRNEEQOREANKKIEKQKQVYRATHRLLLGAGESGKSTIVKQMLHVNFGNPE 71  
Qy 151 -----EKKQKILDIRKNVKDAIIVTIVSAMSTIIPVPLANPENQFRSDYI 195  
Db 72 GGBEDPQAARSNSGGEKATKVQDIKNLKEAETIVAAMSNLVPVVELANPENQFRVDYI 131  
Qy 196 KSTAPITDFEYSQEFFDHVKLWDEGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYT 255  
Db 132 LSNVNVDPDFPFPEFYEHAKALWEDEGVACRYERSNEYQLIDCAQYFLDKIDVVKQDYV 191  
Qy 256 PTQDILLCRLVLTSGIFETRFQVDKVNFMFVGGQDERRKWKIQCNDVYTAIIYVAAS 315  
Db 192 PSQDILLCRLVLTSGIFETRFQVDKVNFMFVGGQDERRKWKIQCNDVYTAIIYVASS 251  
Qy 316 SYNWVREDNNTNRLRESLDFPESINWNRWLTISILFLNKQDMLAEKVLGKSKIEDY 375  
Db 252 SYNWVREDNNTNRLQEQALNLFKSIWNRWLTISILFLNKQDMLAEKVLGKSKIEDY 311  
Qy 376 FPEYANVTVPEDATPDAGEPKVTRAKFFIRDLFLRISTATGDKHYCYPHFTCAVDTEN 435  
Db 312 FPEFARYTTPEDATPEGEDPRVTRAKYFIRDFELRISTAGDGRHYCYPHSTCAVDTEN 371  
Qy 436 IRRVFNDCRDIIOHMLRQYELL 458  
Db 372 IRRVFNDCRDIIOHMLRQYELL 394

## RESULT 14

## S34421

GTP-binding regulatory protein Gs alpha chain - human  
C;Species: Homo sapiens (man)  
C;Date: 08-Dec-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C;Accession: S34421; S40964  
R;Swaroop, A.; Agarwal, N.; Gruen, J.R.; Bick, D.; Weissman, S.M.  
Nucleic Acids Res. 19, 4723-4729, 1991  
A;Title: Differential expression of novel Gs-alpha signal transduction protein cDNA spec  
A;Reference number: S34421; MUID:91367671; PMID:1716359  
A;Accession: S34421  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-419 <SNA>  
C;Cross-references: UNIPROT:Q14455; UNIPARC:UPI0000073FD3; EMBL:X56009; NID:g31916; PIDN:  
R;Ali, I.U.; Reinhold, W.; Salvador, C.; Aguanno, S.  
Nucleic Acids Res. 20, 4263-4267, 1992  
A;Title: Aberrant splicing of Gs-alpha transcript in transformed human astroglial and gl  
A;Reference number: S40964; MUID:92375734; PMID:1324476  
A;Accession: S40964  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 112-116, 236-242, 263-264, 268-275, 279-286; 306-309 <ALI>  
C;Cross-references: UNIPARC:UPI000017513F; UNIPARC:UPI0000177863; UNIPARC:UPI0000177864;  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain

Query Match 60.88; Score 1460; DB 2; Length 419;  
Best Local Similarity 63.48; Pred. No. 8.8e-86;  
Matches 287; Conservative 41; Mismatches 71; Indels 54; Gaps 5;

Qy 26 PVEDAQPAPALAPVRAAARDARTLLPRGGEG-----SPACARPKADPKKEKRQTEQ 80  
Db 1 PVSSAPRRGHSVA---SAPRSLGRQVAGRRGAALPCSLAPGCGAAGASPCP----- 50  
Qy 81 LSAEERAAKERAARVAKRVSGIDRMLDQKRDLOOQTHRLLLGAGESGKSTIVKQMR 140  
Db 51 -GAGRRRAAGRCCLACECTS-----LTCAGESGKSTIVKQMR 86  
Qy 141 ILHVNFGNPE-----EKKQKILDIRKNVKDAIIVTIVSAMSTIIPVPLAN 185  
Db 87 ILHVNFGNPEGGEDPQAARSNSGGEKATKVQDIKNLKEAETIVAAMSNLVPVVELAN 146  
Qy 186 PENQFRSDYIKSIAPITDFEYSQEFFDHVKLWDEGVKACFERSNEYQLIDCAQYFLER 245  
Db 147 PENQFRVDYILSNVNVDPDFPFPEFYEHAKALWEDEGVACRYERSNEYQLIDCAQYFLDK 206  
Qy 246 IDSLSLVDTPTDQDILLCRLVLTSGIFETRFQVDKVNFMFVGGQDERRKWKIQCNDV 305  
Db 207 IDVIKQADYVPSDQDILLCRLVLTSGIFETRFQVDKVNFMFVGGQDERRKWKIQCNDV 266  
Qy 306 TAIIVVAACSSYNWVREDNNTNRLRESLDFPESINWNRWLTISILFLNKQDMLAEKV 365  
Db 267 TAIIFVSSSSYNWVREDNNTNRLQEQALNLFKSIWNRWLTISILFLNKQDMLAEKV 326  
Qy 366 LAGSKIEDYFPYANVTVPEDATPDAGEPKVTRAKFFIRDLFLRISTATGDKHYCYP 425  
Db 327 LAGSKIEDYFPYARYTTPEDATPEGEDPRVTRAKYFIRDFELRISTAGDGRHYCYP 386  
Qy 426 HFTCAVDTENIRRVFNDCRDIIOHMLRQYELL 458  
Db 387 HFTCAVDTENIRRVFNDCRDIIOHMLRQYELL 419

## RESULT 15

## S27015

GTP-binding regulatory protein Gs alpha chain - great pond snail  
N;Alternate names: guanine nucleotide-binding protein Gs alpha-S chain  
C;Species: Lymnaea stagnalis (great pond snail)  
C;Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 09-Jul-2004  
C;Accession: S27015; S25590  
R;Knol, J.C.; Weidemann, W.; Planta, R.J.; Vreugdenhil, E.; van Heerikhuizen, H.  
FEBS Lett. 314, 215-219, 1992

A;Title: Molecular cloning of G protein alpha subunits from the central nervous system  
A;Reference number: S27013; MUID:93106153; PMID:1468550  
A;Accession: S27015  
A;Molecule type: mRNA  
A;Residues: 1-376 <KNO1>  
A;Cross-references: UNIPROT:P30684; UNIPARC:UIP0000177868; EMBL:Z15096  
R;Knol, J.C.; Weidemann, W.; Plantica, R.J.; Vreugdenhil, E.; van Heerikhuizen, H.  
submitted to the EMBL Data Library, September 1992  
A;Description: Molecular cloning of G protein alpha subunits that are expressed in the  
A;Reference number: S25588  
A;Accession: S25590  
A;Molecule type: mRNA  
A;Residues: 1-64, F', 66-93, 'T', 95-358, 'D', 360-376 <KNO2>  
A;Cross-references: UNIPARC:UIP0000128207; EMBL:Z15096; NID:G9634; PIDN:CAA78808.1; PID:  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
C;Keywords: GTP binding; nucleotide binding; P-loop; signal transduction  
F;44-51/Region: nucleotide-binding motif A (P-loop)  
F;274-277/Region: GTP-binding NKXD motif

Query Match	58.9%	Score	1413	DB 2	Length	376			
Best Local Similarity	71.0%	Pred. No.	7.7e-83						
Matches	262	Conservative	50	Mismatches	57	Indels	0	Gaps	0

  

Qy	90	KEREAVEAKRVSGIDMLRDQXRDQLQOTHRLLLLGAGESGKSTIVKOMRILHVGNSP	149
Db	8	RDEDDKLRKEANKKIEKQLAKDKLLYRGTHRLLLLGAGESGKSTIVKOMRILHVGNSP	67
Qy	150	EEKQKILDIRKNVKAIDVITVSAMSTIIPVPLPLANPQRSDYIKSIAPITPFYSQE	209
Db	68	EEKQKIEDIRKNVRDAITLTIGAMSNLNPVQLEHPQNKAKVDYIQDKASQAEFDYPI	127
Qy	210	FFDHVKLWDDEGKVKAFERSNEYQLIDCAQYFLERIDSVLSVDYPTDOLLRCRVLT	269
Db	128	FYEYTEILWKDKGVQAAPERSNEYQLIDCAQYFLDRVHIIRQAEYTPSEQDILRCRVLT	187
Qy	270	GIFETRFQVDKVNFMFDVGQDERRRKWICQFNDVTAIIVAAAGSSNMVIRENNNTR	329
Db	188	GIFETKSVDKVNFMFDVGQDERRRKWICQFNDVTAIIVTACSGYNMVLREDATQNR	247
Qy	330	LRESLDLPESIWNNRWLRTISIIILFNKQDMLAEKVLGAGSKIEDYFPFEYANYTVPEDAT	389
Db	248	LKESLDLPFSIWNNRWLRTISVILFNKQDLAEKVKAGSKIEDYFPFEYARYQVPDPAS	307
Qy	390	PDAGEDPKVTRAKFFIRDLFLRISTATGDGKHGYCPHFTCAVDTENIRRVFNDCRDIIQR	449
Db	308	SEPEDTEVVRKAFYIRDEFRLRISTASGDGRHYCYPHFTCAVDTENIRRVFNDCRDIIQR	367
Qy	450	MHLKQYELL	458
Db	368	MHLRQYELL	376

Search completed: March 2, 2006, 19:31:05  
Job time : 28.1616 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:29:50 ; Search time 286.601 Seconds  
(without alignments)  
1127.462 Million cell updates/sec

Title: US-10-618-320A-1

Perfect score: 2400

Sequence: 1 MGICYSLRPLFGPGDDPC.....VFNDCRDIQRMHLKQYELL 458

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2400	100.0	458	2	Q86XU3 HUMAN
2	2124	88.5	448	2	Q66147 MOUSE
3	1847	77.0	351	2	Q8N2B4 HUMAN
4	1819	75.8	381	1	GNAL HUMAN
5	1819	75.8	381	1	GNAL MOUSE
6	1811	75.5	381	1	GNAL_RAT
7	1755	73.1	462	2	Q6GLL1 XENLA
8	1677	69.9	379	2	Q90WN6 XENLA
9	1639	68.3	379	2	Q6PUS6 BRARE
10	1620	67.5	399	2	Q6DG15 BRARE
11	1609	67.0	379	2	Q5U3H6 BRARE
12	1567.5	65.3	485	2	Q5FNY2 HUMAN
13	1563	65.1	722	2	Q5JWF1 HUMAN
14	1559	65.0	736	2	Q5JWF2 HUMAN
15	1559	65.0	756	2	Q9ZLN8 MOUSE
16	1555	64.8	1133	2	Q6R0H7 MOUSE
17	1548	64.5	715	2	Q63803 RAT
18	1540	64.2	379	2	Q6NXS0 HUMAN
19	1540	64.2	379	2	Q14433 HUMAN
20	1529.5	63.7	380	2	Q5JWD2 HUMAN
21	1528	63.7	379	2	Q6P413 XENLA
22	1527	63.6	379	1	GNAS XENLA
23	1525.5	63.6	394	1	GNAS_CANPA
24	1525.5	63.6	394	1	GNAS HUMAN
25	1525.5	63.6	394	2	Q5JWD5 HUMAN
26	1525	63.5	395	2	Q5JWD4 HUMAN
27	1522.5	63.4	394	1	GNAS BOVIN
28	1521.5	63.4	394	1	GNAS CRIGR
29	1520.5	63.4	394	1	GNAS_CRIL0
30	1520.5	63.4	394	1	GNAS_MESAU
31	1520.5	63.4	394	1	GNAS_MOUSE

32	1520.5	63.4	394	1	GNAS_RAT
33	1520.5	63.4	394	2	Q58E62 MOUSE
34	1519	63.3	379	2	Q6NVN4 XENTR
35	1518	63.2	397	1	GNAS_PIG
36	1507	62.8	342	2	Q8TBC0 HUMAN
37	1495	62.3	340	2	Q6X973 RABIT
38	1471	61.3	379	2	Q6QM14_LYTVA
39	1463	61.0	379	2	Q6QM15_STRPU
40	1462.5	60.9	379	1	GNAS_HOMAM
41	1460	60.8	419	2	Q14455 HUMAN
42	1458.5	60.8	362	2	Q9Z1R8_RAT
43	1457.5	60.7	362	2	Q9Z1R7_MOUSE
44	1457.5	60.7	417	2	Q5JWC5_HUMAN
45	1440.5	60.0	379	2	Q66Q60_PENVA

## ALIGNMENTS

### RESULT 1

ID	Q86XU3_HUMAN PRELIMINARY;	PRT;	458 AA.
AC	Q86XU3;		
DT	01-JUN-2003 (TrEMBLrel. 24, Created)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	Guanine nucleotide binding protein (G protein), alpha activating		
DE	activity polypeptide, olfactory type, isoform 1.		
GN	Names=GNAL;		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]_TaxID=9606;		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Testis;		
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Whiting M., Madan A., Kettman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield J.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,		
RA	Schmurch A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Testis;		
RG	NIH MGC Project;		
RL	Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC050021; AH50021.1; -; mRNA.		
DR	HSSP; P04896; 1CJU.		
DR	SMR; Q86XU3; 117-455.		
DR	Ensembl; ENSG00000141404; Homo sapiens.		
DR	GO; GO:0005525; F:GTP binding; IEA.		
DR	GO; GO:0004871; F:signal transducer activity; IEA.		
DR	GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.		
DR	GO; GO:0007165; P:signal transduction; IEA.		
DR	InterPro; IPR001019; Gprotein_alpha_b.		
DR	InterPro; IPR000367; Gprotein_alpha_s.		

P63095	rattus norv
Q58e62	mus musculu
Q6NVN4	kenopus tro
P29797	sus scrofa
Q8Tbc0	homo sapien
Q6x973	oryctolagus
Q6qm14	lytechinus
Q6qm15	strongyloce
Q16118	homo sapien
Q14455	homo sapien
Q9Z1R8	rattus norv
Q9Z1R7	mus musculu
Q5Jwc5	homo sapien
Q66q60	penaeus van

```

DR InterPro; IPR011025; Gproteina_insert.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR PRINTS; PR00443; GPROTEINAS.
DR PRODOM; PD000281; Gproteina_alpha; 1.
DR SMART; SM00275; G_alpha; 1.
SQ SEQUENCE 458 AA; 52455 MW; 89231879924C06A1 CRC64;

Query Match 100.0%; Score 2400; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 7.9e-135;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLCYSLRPLFLFGPGDDPCAASEPPVEDAQAPAPALAPVRAAARDTARTLLPRGEGS 60
DB 1 MGLCYSLRPLFLFGPGDDPCAASEPPVEDAQAPAPALAPVRAAARDTARTLLPRGEGS 60
QY 61 PACARPKADPKPKRQRTQOLSABEREAAKREAVKARVSRGIDRMRLDQKRDLDQOOTH 120
DB 61 PACARPKADPKPKRQRTQOLSABEREAAKREAVKARVSRGIDRMRLDQKRDLDQOOTH 120
QY 121 RLLLLGAGESGKSTIVKQMRILHVNGFNPEKKQKILDIRKNVKDAIVTIVSAMSTIIPP 180
DB 121 RLLLLGAGESGKSTIVKQMRILHVNGFNPEKKQKILDIRKNVKDAIVTIVSAMSTIIPP 180
QY 181 VPLANPENQFRSDYIKSIAPITDFEYSQEFFDHHVKKLWDEGVKACFERSNEYQLIDCAQ 240
DB 181 VPLANPENQFRSDYIKSIAPITDFEYSQEFFDHHVKKLWDEGVKACFERSNEYQLIDCAQ 240
QY 241 YFLERIDSVSLVDYPTPDQDLRCRVLTSGIFETRFOVDKVNFMFVGVGQQRDERKWKIQ 300
DB 241 YFLERIDSVSLVDYPTPDQDLRCRVLTSGIFETRFOVDKVNFMFVGVGQQRDERKWKIQ 300
QY 301 CFNDVTAIIVAACSSYNMVRDNNNTNRLRESLDLFESEIWNRLWRTISILFLNKQDM 360
DB 301 CFNDVTAIIVAACSSYNMVRDNNNTNRLRESLDLFESEIWNRLWRTISILFLNKQDM 360
QY 361 LAEVLGAGSKIEDYFPEYANYTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK 420
DB 361 LAEVLGAGSKIEDYFPEYANYTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK 420
QY 421 HYCPHPFTCAVDTENIRRVNDCRDIIRMHKQYELL 458
DB 421 HYCPHPFTCAVDTENIRRVNDCRDIIRMHKQYELL 458

RESULT 2
Q66L47_MOUSE
ID Q66L47_MOUSE PRELIMINARY; PRT; 448 AA.
AC Q66L47;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Gnal protein.
GN Gnal protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

Query Match 88.5%; Score 2124; DB 2; Length 448;
Best Local Similarity 91.0%; Pred. No. 2.1e-118;
Matches 417; Conservative 6; Mismatches 25; Indels 10; Gaps 3;

QY 1 MGLCYSLRPLFLFGPGDDPCAASEPPVEDAQAPAPALAPVRAAARDTARTLLPRGEGS 60
DB 1 MGLCYSLRPLFLFGSPEDTCAASEPCAEDAQPSAAPASIPAPA--PVGTLRRGGRI 58
QY 61 PACARPKADPKPKRQRTQOLSABEREAAKREAVKARVSRGIDRMRLDQKRDLDQOOTH 120
DB 59 VANARPPGE--LQSRRRQQLRAEEREA-----KEARKVSRGIDRMRLREQKRDLDQOOTH 110
QY 121 RLLLLGAGESGKSTIVKQMRILHVNGFNPEKKQKILDIRKNVKDAIVTIVSAMSTIIPP 180
DB 111 RLLLLGAGESGKSTIVKQMRILHVNGFNPEKKQKILDIRKNVKDAIVTIVSAMSTIIPP 170
QY 181 VPLANPENQFRSDYIKSIAPITDFEYSQEFFDHHVKKLWDEGVKACFERSNEYQLIDCAQ 240
DB 171 VPLANPENQFRSDYIKSIAPITDFEYSQEFFDHHVKKLWDEGVKACFERSNEYQLIDCAQ 230
QY 241 YFLERIDSVSLVDYPTPDQDLRCRVLTSGIFETRFOVDKVNFMFVGVGQQRDERKWKIQ 300
DB 231 YFLERIDSVSLVDYPTPDQDLRCRVLTSGIFETRFOVDKVNFMFVGVGQQRDERKWKIQ 290
QY 301 CFNDVTAIIVAACSSYNMVRDNNNTNRLRESLDLFESEIWNRLWRTISILFLNKQDM 360
DB 291 CFNDVTAIIVAACSSYNMVRDNNNTNRLRESLDLFESEIWNRLWRTISILFLNKQDM 350
QY 361 LAEVLGAGSKIEDYFPEYANYTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK 420
DB 351 LAEVLGAGSKIEDYFPEYANYTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK 410
QY 421 HYCPHPFTCAVDTENIRRVNDCRDIIRMHKQYELL 458
DB 411 HYCPHPFTCAVDTENIRRVNDCRDIIRMHKQYELL 448

RESULT 3
Q8N2B4_HUMAN
ID Q8N2B4_HUMAN PRELIMINARY; PRT; 351 AA.
AC Q8N2B4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)

```

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein FLJ33549.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Amalgam;  
 RX PubMed=14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J.-I., Saito K., Kawai I., Isono Y., Nakamura Y.,  
 RA Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,  
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,  
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,  
 RA Omura Y., Abe K., Kamiyama K., Katsuta N., Sato K., Tanikawa M.,  
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,  
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,  
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotuta T.,  
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,  
 RA Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
 RA Yoshihara Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kunagai A., Itakura S., Fukuzumi Y.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Kitatani R., Kakajima T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.";  
 RL Nat. Genet. 36:40-45(2004).  
 DR EMBL; AK090868; BAC03935.1; -- mRNA.  
 DR HSP; P04896; 1CJU.  
 DR SMR; Q8N2B4; 10-348.  
 DR GO; GO:0005525; F:GTP binding; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR011025; Gprotein\_alpha\_insert.  
 DR InterPro; IPR001019; Gprotein\_alpha.  
 DR InterPro; IPR000367; Gprotein\_alpha\_s.  
 DR Pfam; PF00503; G-alpha\_1.  
 DR PRINTS; PR03148; GPROTEIN.  
 DR PRINTS; PR00443; GPROTEINAS.  
 DR ProDom; PD000281; Gprotein\_alpha: 1.  
 DR SMART; SM00275; G\_alpha: 1.  
 SQ SEQUENCE 351 AA; 41057 MW; 4956B9D7573F9F60 CRC64;  
 Query Match 77.0%; Score 1847; DB 2; Length 351;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-102;  
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 108 MLRDKEDLQTHLLGAGESKSTIVKQMRILHVGNSPEKKKILDIRNKVDAL 167  
 DB 1 MLRDKEDLQTHLLGAGESKSTIVKQMRILHVGNSPEKKKILDIRNKVDAL 60  
 QY 168 VTVSAMSTIIPVPLANPENFRSDYIKSIAPITDPYEQSFDFHVKLWDDEGVKACF 227  
 DB 61 VTVSAMSTIIPVPLANPENFRSDYIKSIAPITDPYEQSFDFHVKLWDDEGVKACF 120  
 QY 228 ERSEYQLIDCAQYFLERIDSVLVTPTDQDLLRCRVLTSIGIFTRFQVDKVNFMFD 287

Db 121 ERSEYQLIDCAQYFLERIDSVLVTPTDQDLLRCRVLTSIGIFTRFQVDKVNFMFD 180  
 QY 288 VGGQDERRKWIQCFNDVTATIIYVACSSYNMVIREDNNTNRLRESLDLFESINWNL 347  
 Db 181 VGGQDERRKWIQCFNDVTATIIYVACSSYNMVIREDNNTNRLRESLDLFESINWNL 240  
 QY 348 TISIIILFLNKLQDLAEKVLGKSKIEDYFPEYANYTVPEDATPDAGDPKVKTRAKFIRD 407  
 Db 241 TISIIILFLNKLQDLAEKVLGKSKIEDYFPEYANYTVPEDATPDAGDPKVKTRAKFIRD 300  
 QY 408 LFLRISTATGDKGHYCPHFTCAVDENIRRVFNDRCRDIQRMHLKQYELL 458  
 Db 301 LFLRISTATGDKGHYCPHFTCAVDENIRRVFNDRCRDIQRMHLKQYELL 351

RESULT 4  
 ID GNAL HUMAN STANDARD; PRT; 381 AA.  
 AC P38405;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 48, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Guanine nucleotide-binding protein G(olf), alpha subunit (Adenylate  
 DE cyclase-stimulating G alpha protein, olfactory type).  
 DE Name=GNAL;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Insulinoma;  
 RX MEDLINE=94062616; PubMed=8243272; DOI=10.1210/en.133.6.2508;  
 RA Zisman J.M.; Westmark G.T.; Lamendola J.; Boel E.; Steiner D.F.;  
 RT "Human G(olf) alpha: complementary deoxyribonucleic acid structure and  
 RT expression in pancreatic islets and other tissues outside the  
 RT olfactory neuroepithelium and central nervous system.";  
 RL Endocrinology 133:2508-2514(1993).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Vuorio J.T., Berrettini W.H., Overhauser J., Prockop D.J.,  
 RA Ferraro T.N., Ala-Kokko L.;  
 RT "The gene for the human G protein Golf alpha.";  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Brain;  
 RA Puhl H.L. III, Ikeda S.R., Aronstam R.S.;  
 RT "cDNA clones of human proteins involved in signal transduction  
 RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are  
 CC involved as modulators or transducers in various transmembrane  
 CC signaling systems. G(olf) alpha mediates signal transduction  
 CC within the olfactory neuroepithelium and the basal ganglia. May be  
 CC involved in some aspect of visual transduction, and in mediating  
 CC the effect of one or more hormones/neurotransmitters.  
 CC -1- SUBUNIT: G proteins are composed of 3 units; alpha, beta and  
 CC gamma. The alpha chain contains the guanine nucleotide binding  
 CC site.  
 CC -1- TISSUE SPECIFICITY: Detected in olfactory neuroepithelium, brain,  
 CC testis, and to a lower extent in retina, lung alveoli, spleen.  
 CC Trace amounts where seen in kidney, adrenal gland and liver. Found  
 CC to be expressed in all the insulinomas examined.  
 CC -1- SIMILARITY: Belongs to the G-alpha family. G(s) subfamily.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.

```

CC EMBL; L10665; AAC37535.1; -; mRNA.
DR EMBL; U55184; AAD00085.1; -; Genomic DNA.
DR EMBL; U55180; AAD00085.1; JOINED; Genomic DNA.
DR EMBL; U55181; AAD00085.1; JOINED; Genomic DNA.
DR EMBL; U55182; AAD00085.1; JOINED; Genomic DNA.
DR EMBL; U55183; AAD00085.1; JOINED; Genomic DNA.
DR EMBL; AF493893; AAM12607.1; -; mRNA.
DR PIR; I53271; I53271.
DR HSSP; P04896; 1CJU.
DR SMR; P38405; 39-378.
DR ENSEMBL; ENSG0000041404; Homo sapiens.
DR HGNC; HGNC:4388; GNAL.
DR MIM; 139312; -.
DR GO; GO:0003294; F:GTPase activity; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR001019; Gprotein_alpha_bd.
DR InterPro; IPR000367; Gprotein_alpha_S.
DR InterPro; IPR011025; GproteinA_insert.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR PRINTS; PR00443; GPROTEINAS.
DR ProDom; PD000281; Gprotein_alpha; 1.
KW ADP-ribosylation; GTP-binding; Lipoprotein; Multigene family;
KW Nucleotide-binding; Palmitate; Transducer.
FT NP_BIND 49 56 GTP (By similarity).
FT NP_BIND 210 214 GTP (By similarity).
FT NP_BIND 279 282 GTP (By similarity).
FT MOD_RES 188 188 ADP-ribosylarginine (by cholera toxin).
FT LIPID 3 3 S-palmitoyl cysteine (By similarity).
SQ SEQUENCE 381 AA; 44308 MW; 9A73D2982FF9C9A3 CRC64;

Query Match 75.8%; Score 1819; DB 1; Length 381;
Best Local Similarity 95.0%; Pred. No. 2.5e-100;
Matches 345; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

Qy 96 KEARKVSRGIDRLMDQKRLQDQLQTHRLLLGAGESGKSTIVKQWRIHVNFGFPEKKQK 155
Db 19 KERREANKIEKQLQKERLAYKATHRLLLGAGESGKSTIVKQWRIHVNFGFPEKKQK 78

Qy 156 ILDIRKNVKAIVTVISAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQEFDFHVK 215
Db 79 ILDIRKNVKAIVTVISAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQEFDFHVK 138

Qy 216 KLWDDGVKACFRSNEYQLIDCAQYELERIDSVLDVYPTDQDLRLCHVLTSGIFETR 275
Db 139 KLWDDGVKACFRSNEYQLIDCAQYELERIDSVLDVYPTDQDLRLCHVLTSGIFETR 198

Qy 276 FQVDKVNFMFDVGGQDERRKWTQCFNDVTAIYVAACSSYNVIREDNNTNRLRESLD 335
Db 199 FQVDKVNFMFDVGGQDERRKWTQCFNDVTAIYVAACSSYNVIREDNNTNRLRESLD 258

Qy 336 LFESIMNRLWRTISILFLNKQDMLAEKVLAKGSKIEDYFPEYANYTPEDATPDAGED 395
Db 259 LFESIMNRLWRTISILFLNKQDMLAEKVLAKGSKIEDYFPEYANYTPEDATPDAGED 318

Qy 396 PKYTRAKFFRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFNDCRDIIRMHLKQY 455
Db 319 PKYTRAKFFRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFNDCRDIIRMHLKQY 378

Qy 456 ELL 458
Db 379 ELL 381

RESULT 5
ID GNAL_MOUSE STANDARD; PRT; 381 AA.
AC QRCGK7; Q61020; Q61589;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Guanine nucleotide-binding protein G(olf), alpha subunit (Adenylate
cyclase-stimulating G alpha protein, olfactory type).
Name=Gnal;
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
Muridae; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE.
STRAIN=C57BL/6;
Von Dannecker L.E.C., Malnic B.;
"The mouse G protein Golf alpha subunit full length coding sequence.";
Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE OF 209-373.
STRAIN=CF-1 / Harlan; TISSUE=Brain;
MEDLINE=97011591; PubMed=8858601;
DOI=10.1002/(SICI)1098-2795(199607)44:3<315::AID-MRDS-3.3.CO;2-V;
Williams C.J., Schultz R.M., Kopf G.S.;
"G protein gene expression during mouse oocyte growth and maturation,
and preimplantation embryo development.";
Mol. Reprod. Dev. 44:315-323(1996).
[3]
RN NUCLEOTIDE SEQUENCE OF 216-276.
TISSUE=Brain;
MEDLINE=90017488; PubMed=2508088;
Strathmann M., Wilkie T.M., Simon M.I.;
"Diversity of the G-protein family: sequences from five additional
alpha subunits in the mouse.";
Proc. Natl. Acad. Sci. U.S.A. 86:7407-7409(1989).
CC -!- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are
involved as modulators or transducers in various transmembrane
signaling systems. G(olf) alpha mediates signal transduction
within the olfactory neuroepithelium and the basal ganglia. May be
involved in some aspect of visual transduction, and in mediating
the effect of one or more hormones/neurotransmitters (By
similarity).
CC -!- SUBUNIT: G proteins are composed of 3 units; alpha, beta and
gamma. The alpha chain contains the guanine nucleotide binding
site.
CC -!- SIMILARITY: Belongs to the G-alpha family. G(s) subfamily.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; AV179169; AAC03564.1; -; mRNA.
CC EMBL; U38503; AAB01734.1; -; mRNA.
CC EMBL; M57635; AAG63300.1; -; mRNA.
CC HSSP; P04896; 1CJU.
CC SMR; QRCGK7; 39-378.
CC ENSEMBL; ENSMUSG00000024524; Mus musculus.
CC MGI; MGI:95774; Gnal.
CC GO; GO:0007190; P:adenylate cyclase activation; IDA.
CC GO; GO:0007608; P:perception of smell; IMP.
CC InterPro; IPR001019; Gprotein_alpha_bd.
CC InterPro; IPR000367; Gprotein_alpha_S.
CC InterPro; IPR011025; GproteinA_insert.
CC Pfam; PF00503; G-alpha; 1.
CC PRINTS; PR00318; GPROTEINA.
CC PRINTS; PR00443; GPROTEINAS.
CC ProDom; PD000281; Gprotein_alpha; 1.
CC SMART; SM00275; G_alpha; 1.
KW ADP-ribosylation; GTP-binding; Lipoprotein; Multigene family;
KW Nucleotide-binding; Palmitate; Transducer.
FT NP_BIND 49 56 GTP (By similarity).
FT NP_BIND 210 214 GTP (By similarity).
FT NP_BIND 279 282 GTP (By similarity).
FT MOD_RES 188 188 ADP-ribosylarginine (by cholera toxin).
FT LIPID 3 3 S-palmitoyl cysteine (By similarity).

```





ID Q6GLL1\_XENLA PRELIMINARY; PRT; 462 AA.  
 AC Q6GLL1, 2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE LOC43714 protein (fragment).  
 GN Name=LOC43714;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Tohiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,  
 RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative."  
 RL Dev. Dyn. 225:394-391(2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC074466; AAH74466.1; -, mRNA.  
 DR SRR; Q6GLL1; 120-459.  
 DR GO; GO:0005525; F:GTP binding; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0007186; P:g-protein coupled receptor protein signalin. . .; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR011025; GproteinA\_insert.  
 DR InterPro; IPR001019; Gprotein\_alpha.  
 DR InterPro; IPR00367; Gprotein\_alpha\_S.  
 DR Pfam; PF00503; G-alpha; 1.  
 DR PRINTS; PR00318; GPROTEINA.  
 DR PRODOM; PD000281; Gprotein\_alpha; 1.  
 DR SMART; SM00275; G\_alpha; 1.  
 DR TIGRFAMs; TIGR00756; PPR; 1.  
 FT NON TER 1  
 SQ SEQUENCE 462 AA; 53392 MW; 57EB03C65C17DAF4 CRC64;  
 Query Match 73.4%; Score 1755; DB 2; Length 462;  
 Best Local Similarity 74.5%; Pred. No. 2.le-96;  
 Matches 345; Conservative 32; Mismatches 56; Indels 30; Gaps 4;

QY 15 PGDPCAASEPVPVEDAQAPAPALAPVRAA-ARDTARTLLPRGEGSPACARPKADK--- 70  
 DB 11 PG-----ASOTGDVPCQPRRVEQCGRGGDSADGLVLLQNGGGSGESQKHKASKAPH 65  
 QY 71 -----PKEKQRTEQLSABERAAKEREAVKARKVSRGIDRLMRQDKRD 115  
 DB 66 POLQHPHPQHPPPPLHKGWDLRLKEAE-----KEAKVKSTIDRLVLEQKRE 119  
 QY 116 LOOTHRLLLLGAGESGKSTIVKQMRLLHVNGNPPEKKOKILDIRKNVKAIVTIVSAMS 175  
 DB 120 YKOTHRLLLLGAGESGKSTIVKQMRLLHVNGNPSEKKOKIQDIRKNVKAIVTIVSAMS 179  
 QY 176 TIIPVPLANPENQFRSDYIKSIAPITDPESYQEFDFHVKKLWDDGKVCACFERSNEYQL 235  
 DB 180 TIIPVPLANPENQFRSLDYIKSIAPLSDFDYIQEFHQAQKLWDDGKVCACFERSNEYQL 239  
 QY 236 IDCAQYFLERIDSVLSVDVTPDQDLLRCRLVTLSGIFETFRQVDKVNFMFVGGQORDER 295  
 DB 240 IDCAQYFLERIDVRQNDVTPDQDLLRCRLVTLSGIFETFRQVDKVNFMFVGGQORDER 299  
 QY 296 RKWIOCFNDVTAILIYVACSSYNMVIREDNNTNRLRESLDLFSIWNNEWLRITISILFEL 355  
 DB 300 RKWIOCFNDVTAILIYVACSSYNMVIREDNNTNRLREALDLFKSIWNNEWLRITISILFEL 359  
 QY 356 NKQDMLAEKVLAKGSKIEDYFPEYANYTVPEDATPDAGDPKVTAKFFIRDLFRISTA 415  
 DB 360 NKQDMLAEKVLAKGSKIEDYFPEYVYTIPTDAIPDAGDPKVTAKFFIRDEFIRISTA 419  
 QY 416 TGDGRHYCPHFTCAVDTEINRRVNDCRDIITORMHLKQYELL 458  
 DB 420 SGDGRHYCPHFTCAVDTEINRRVNDCRDIITORMHLRQYELL 462  
 RESULT 8  
 Q90WN6\_XENLA PRELIMINARY; PRT; 379 AA.  
 ID Q90WN6  
 AC Q90WN6;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE G-protein alpha subunit.  
 GN Name=g(alpha)ex2;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=21479556; PubMed=11596062; DOI=10.1002/cne.1358;  
 RA Mezler M., Fleischer J., Conzelmann S., Korch A., Widmayer P.,  
 RA Breer H., Boekhoff I.;  
 RT "Identification of a nonmammalian G(olf) subtype: Functional role in  
 RT olfactory signaling of airborne odors in Xenopus laevis."  
 RL J. Comp. Neurol. 439:400-410(2001).  
 DR EMBL; AJ296281; CAC82735.1; -, mRNA.  
 DR HSSP; P04896; 1CJU.  
 DR SMR; Q90WN6; 37-376.  
 DR GO; GO:0005525; F:GTP binding; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0007186; P:g-protein coupled receptor protein signalin. . .; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR011025; GproteinA\_insert.  
 DR InterPro; IPR001019; Gprotein\_alpha.  
 DR InterPro; IPR00367; Gprotein\_alpha\_S.  
 DR Pfam; PF00503; G-alpha; 1.  
 DR PRINTS; PR00318; GPROTEINA.  
 DR PRODOM; PD000281; Gprotein\_alpha; 1.  
 DR SMART; SM00275; G\_alpha; 1.  
 DR TIGRFAMs; TIGR00756; PPR; 1.  
 SQ SEQUENCE 379 AA; 44366 MW; 06D2D8E1C8918284 CRC64;

```
Query Match 69.9%; Score 1677; DB 2; Length 379;
Best Local Similarity 86.2%; Pred. No. 7.2e-92;
Matches 313; Conservative 25; Mismatches 25; Indels 0; Gaps 0;

QY 96 KEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGSGKSTIVKQWRIHLVNGFNPPEKKQK 155
DB 17 KAQREANKKIEKQLOKEREYATKATHRLLLGAGSGKSTIVKQWRIHLVNGFNPPEKKQK 76

QY 156 ILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQBFPHVK 215
DB 77 SQDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQBFPHVK 136

QY 216 KLWDDGVKACFERSNEYQLIDCAQYFLERIDSVLVDTPTDQDLRCRVLTSIGIFETR 275
DB 137 KLWDDGVKACFERSNEYQLIDCAQYFLERIDSVLVDTPTDQDLRCRVLTSIGIFETR 196

QY 276 FOVDKVNFMFDVGGQDERRKWIQCFNDVTALIIYVAACSSYNNVIREDNNTNRLRESLD 335
DB 197 FOVDKVNFMFDVGGQDERRKWIQCFNDVTALIIYVAACSSYNNVIREDNNTNRLRESLD 256

QY 336 LFESINNNRRLRTISIIILFNKQDMLAEKVLGKSKIEDYFPEYVNTVPEDATPDAGED 395
DB 257 LFESINNNRRLRTISIIILFNKQDMLAEKVLGKSKIEDYFPEYVNTVPEDATPDAGED 316

QY 396 PKVTRAKFFIRDLFLRISTATGDKHKYCYHFTCAVDENIRRVFNDCRDIIQRMHLKQY 455
DB 317 PKVTRAKFFIRDLFLRISTATGDKHKYCYHFTCAVDENIRRVFNDCRDIIQRMHLKQY 376

QY 456 ELL 458
DB 377 ELL 379

RESULT 9
Q6PUS6 BRARE
ID Q6PUS6 BRARE PRELIMINARY; PRT; 379 AA.
AC Q6PUS6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE GNAS complex locus.
GN ORFNames=zgc:92392;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhang G.W., Sun X.J., Wu X.Y., Song H.D., Liu T.X., Zhou Y., Sheng Y.,
RA Chen Y., Ruan Z., Jiang C.L., Fan H.Y.; Zon L.I., Kanki J.P.,
RA Look A.T., Chen Z.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY576989; AAS92627.1; -; mRNA.
DR HSP; P04896; IAZS.
DR SMR; Q6PUS6; 35-376.
DR ZFIN; ZDB-GENE-040718-441; zgc:92392.
DR GO; GO:0005525; P:GTP binding; IEA.
DR GO; GO:0004871; P:signal transducer activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR011025; GProteinA insert.
DR InterPro; IPR001019; GProtein_alpha.
DR InterPro; IPR000367; GProtein_alpha_S.
DR InterPro; IPR002885; PPR.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEIN.
DR PRINTS; PR00443; GPROTEIN.
DR ProDom; PD000281; GProtein_alpha; 1.
DR SMART; SM00275; G_alpha; 1.
DR TIGRFAH; TIGR00756; PPR; 1.
SEQUENCE 379 AA; 44376 MW; EF81C5DB6D463449 CRC64;
```

```

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC076540; AAH76540.1; -; mRNA.
DR SMR; Q6DGL5; 35-396.
DR ZFIN; ZDB-GENE-040718-441; zgc:92392.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0004871; P:signal transducer activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR011025; G-protein_alpha.1.
DR InterPro; IPR001019; Gprotein_alpha_insert.
DR InterPro; IPR00367; Gprotein_alpha_s.
DR InterPro; IPR002885; pPR.
DR Pfam; PF00503; G-alpha.1.
DR PRINTS; PR00318; GPROTEINAS.
DR PRINTS; PR00443; GPROTEINAS.
DR ProDom; PD000281; Gprotein_alpha; 1.
DR SMART; SM00275; G_alpha; 1.
DR TIGRFAMs; TIGR00756; pPR; 1.
SQ SEQUENCE 399 AA; 46859 MW; AELAGAB67D6FCAF7 CRC64;

Query Match 67.5%; Score 1620; DB 2; Length 399;
Best Local Similarity 79.1%; Pred. No. 1.9e-88;
Matches 303; Conservative 35; Mismatches 25; Indels 20; Gaps 1;

QY 96 KEARKVSRGIDRLMDQKRDLOQTHRLLLLGAGESGKSTIVKQMRILHVNFGNPEKKQK 155
DB 17 KAQREANKKIEKQLOKQERQAYKATHRLLLLGAGESGKSTIVKQMRILHVNFGNPEKKQK 76

QY 156 ILDIRKNVKDAIVTVSAMSTIIPVPLANPENQFRSDYIKSIAPITDPFYSOEFFDHVK 215
DB 77 IQDIRKNVKDAIVTVSAMSTIIPVPLANPEQOFRIDYIKSIAPISDFYTOEFDHAK 136

QY 216 KLWDDGVKACFERSNEYQLIDCAQYFLERIDSVLSVDYITPTDQDLLRCRVLTSIGIFETR 275
DB 137 KLWDEGVKACYERSNEYQLIDCAQYFLERIDAVRQSDYITPTDQDLLRCRVLTSIGIFETR 196

QY 276 FOVDKVNFMFDVGGQDERRKWIQCFNDVTAIYVAACSSYNMVIREDNNTNRLRESLD 335
DB 197 FOVDKVNFMFDVGGQDERRKWIQCFNDVTAIYVAACSSYNMVIREDNNTNRLRESLD 256

QY 336 LFESIWNNRWLTISILFLNKQDMLAEKVLGKSKIEDYFPEYANYTPEATPDGAG 395
DB 257 LFRSIWNNRWLTISILFLNKQDMLAEKVLGKSKIEDYFPEYANYTPEATPDGAG 316

QY 396 PKVTRAKFFIRDLFLRISTATGGHYCYPHETCAVDTENIRRVNDCRDIQRMHLKQY 455
DB 317 PKVTRAKFFIRDLFLRISTATGGHYCYPHETCAVDTENIRRVNDCRDIQRMHLKQY 376

QY 456 ELL 458
|||

Query Match 67.0%; Score 1609; DB 2; Length 379;
Best Local Similarity 82.6%; Pred. No. 8e-88;
Matches 300; Conservative 34; Mismatches 29; Indels 0; Gaps 0;

QY 96 KEARKVSRGIDRLMDQKRDLOQTHRLLLLGAGESGKSTIVKQMRILHVNFGNPEKKQK 155
DB 17 KAQREANKKIEKQLOKQERQAYKATHRLLLLGAGESGKSTIVKQMRILHVNFGNPEKKQK 76

QY 156 ILDIRKNVKDAIVTVSAMSTIIPVPLANPENQFRSDYIKSIAPITDPFYSOEFFDHVK 215
DB 77 ILDIRKNVKDAIVTVSAMSTIIPVPSIANPSNQRAEYIKSIAPISDFYTOEFEFHA 136

QY 216 KLWDDGVKACFERSNEYQLIDCAQYFLERIDSVLSVDYITPTDQDLLRCRVLTSIGIFETR 275
DB 137 HLWDDGVKACFERSNEYQLIDCAQYFLERISVRQNDYITPTDQDLLRCRVLTSIGIFETR 196

QY 276 FOVDKVNFMFDVGGQDERRKWIQCFNDVTAIYVAACSSYNMVIREDNNTNRLRESLD 335
DB 197 FOVDKVNFMFDVGGQDERRKWIQCFNDVTAIYVAACSSYNMVIREDNNTNRLRESLD 256

QY 336 LFESIWNNRWLTISILFLNKQDMLAEKVLGKSKIEDYFPEYANYTPEATPDGAG 395
DB 257 LFRSIWNNRWLTISILFLNKQDMLAEKVLGKSKIEDYFPEYANYTPEATPDGAG 316

QY 396 PKVTRAKFFIRDLFLRISTATGGHYCYPHETCAVDTENIRRVNDCRDIQRMHLKQY 455
DB 317 PKVTRAKFFIRDLFLRISTATGGHYCYPHETCAVDTENIRRVNDCRDIQRMHLKQY 376

QY 456 ELL 458
|||

```

```

Db 377 ELL 379

RESULT 12
QSFY2_HUMAN
ID QSFY2_HUMAN PRELIMINARY; PRT; 485 AA.
AC QSFY2
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE GNAS protein (Fragment).
GN Name=GNAS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Breast;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Breast;
RA Director JGC Project;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC089157; AAH89157.1; -; mRNA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR011025; Gprotein insert.
DR InterPro; IPR001019; Gprotein_alpha.
DR InterPro; IPR000367; Gprotein_alpha_S.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR PRINTS; PR00443; GPROTEINAS.
DR ProDom; PD000281; Gprotein_alpha; 1.
DR SMART; SM00275; G_alpha; 1.
FT NON_TER
SQ
SEQUENCE 485 AA; 54745 MW; 0CF4E28B9E52A1A5 CRC64;

Query Match
Best Local Similarity 65.3%; Score 1567.5; DB 2; Length 485;
Matches 308; Conservative 48; Mismatches 89; Indels 23; Gaps 5;

QY 8 RPLFGGPDPCAAEPVVEDAQAPALAPVRAAARDARTLL-----PRG 56
Db 24 RPLRGK-----AARAPRSRHPPLPARPRAP-RPAARSPRRALLAEEPSRPARPPGA 77
QY 57 GEGSPACARPKAKPKKPKRTQLSAE-----EREAKEREAVKARKVSGIDRLMRD 111
Db 78 APALPARVPAPAAAAARPPAAAAAAMGCLGNSKTEDQRNEEKAQREANKKIEKLOK 137

```

```

QY 112 QKRDLOOQTHRLLLLGAGSGSKSTIVKQMRILHVNNGNPE-EKKQKILDIRKNVDAIVTI 170
Db 138 DKQVYRATHRLLLLGAGSGSKSTIVKQMRILHVNNGNPESEKATKVQDIDKNLKEAIEI 197
QY 171 VSRMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSQEFDFHVKLWDDGKVCACFERS 230
Db 198 VAMSNLVPVPELANPENQFRVDYILSMVNPDPDFPEFEYEHAKALWEDEGVACFERS 257
QY 231 NEYQLIDCAQYFLERIDSVSLVDYPTDQDLRLCRVLTSIGIFETRFOVDKVNHFMDVGG 290
Db 258 NEYQLIDCAQYFLDKIDVIKQADVPSDQDLRLCRVLTSIGIFETRFOVDKVNHFMDVGG 317
QY 291 ORDERRKWICFNDVTAIIVAAACSSYNMVRDNNNTNRLRESLDLPESVWNRWLRTIS 350
Db 318 ORDERRKWICFNDVTAIIVAAACSSYNMVRDNNNTNRLRESLDLPESVWNRWLRTIS 377
QY 351 IILFLANKQDMLAEKVLGKSKIEDIPEPEYANYTVPEATPDAGEDPKVTRAKYFIRDLFL 410
Db 378 VILFLANKQDMLAEKVLGKSKIEDIPEPEYANYTVPEATPDAGEDPKVTRAKYFIRDLFL 437
QY 411 RISTATGDGRHYCYPHFTCAVDTENIRRVFVENDCRDIIRMHKQYELL 458
Db 438 RISTASGDGRHYCYPHFTCAVDTENIRRVFVENDCRDIIRMHKQYELL 485

RESULT 13
Q5JWF1_HUMAN
ID Q5JWF1_HUMAN PRELIMINARY; PRT; 722 AA.
AC Q5JWF1
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE GNAS complex locus.
GN Name=GNAS; ORFName=RP4-543J19.6-009;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Laird G.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Lloyd D.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Moore M.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL109840; CAI42933.1; -; Genomic DNA.
DR EMBL; AL132655; CAI43074.1; -; Genomic DNA.
DR EMBL; AL121917; CAI42567.1; -; Genomic DNA.
DR EMBL; AL109840; CAI43074.1; JOINED; Genomic DNA.
DR EMBL; AL121917; CAI43074.1; JOINED; Genomic DNA.
DR EMBL; AL121917; CAI42933.1; JOINED; Genomic DNA.
DR EMBL; AL132655; CAI42933.1; JOINED; Genomic DNA.
DR EMBL; AL109840; CAI42567.1; JOINED; Genomic DNA.
DR EMBL; AL132655; CAI42567.1; JOINED; Genomic DNA.
DR SMR; Q5JWF1; 379-719.
DR Ensembl; ENSG0000087460; Homo sapiens.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001019; Gprotein_alpha_bd.
DR InterPro; IPR000367; Gprotein_alpha_S.
DR InterPro; IPR011025; GproteinA_insert.
DR InterPro; IPR002885; PPR.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.

```

DR PRINTS; PR00443; GPROTEINAS.  
 DR ProDom; PD000281; Gprotein\_alpha; 1.  
 DR SMART; SM00275; G\_alpha; 1.  
 DR TIGRFAMs; TIGR00756; PPR; 1.  
 SQ SEQUENCE 722 AA; 78821 MW; EC72702CCB835A90 CRC64;

Query Match 65.1%; Score 1563; DB 2; Length 722;  
 Best Local Similarity 70.7%; Pred. No. 9.5e-85;  
 Matches 299; Conservative 50; Mismatches 68; Indels 6; Gaps 4;

QY 42 RAAARDARTLLPR---GGEGSPACARPKADKP-KEKRQTEQLSAEEREAAKEREAVKE 97  
 DB 300 RKQPNLLRNFLVQAFGCGFGRSESPQKASRLKVKVPLAEKRQMRQKAELEKRAQKR 359  
 QY 98 A-RKVSIGIDRLMDQKEDLQOQTHRLLLLGAGSGSKSTIVKQRIILHVGFNPE-EKKQK 155  
 DB 360 AEKRSKLIDKQLQDEKMGVCTHRLLLLGAGSGSKSTIVKQRIILHVGFNFGDSEKATK 419  
 QY 156 ILDIRKNVKDAIVTVSAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSOEFFDHVK 215  
 DB 420 VQDIKNLKEAETIVAAASNLPVVELANPENQFRVDYILSVNVPDFFPPFEYEHAK 479  
 QY 216 KLMDDEGVKACFRSNEYQLIDCAQYFLERIDSVLSVDYTTDQDLRCRVLTSIGIPETR 275  
 DB 480 ALWEDEGVACFRSNEYQLIDCAQYFLERIDSVLSVDYTTDQDLRCRVLTSIGIPETR 539  
 QY 276 FOVDKYNFHFVDGGQDERRKWIQCFNDVTAIIYVAACSSYNVIREDNNTNRLRESLD 335  
 DB 540 FOVDKYNFHFVDGGQDERRKWIQCFNDVTAIIYVAACSSYNVIREDNNTNRLRESLD 599  
 QY 336 LFSIWNRRWLRTISILFLNKQDLAEKVLGAGSKIEDYFPEYANTVTPEDATPDAGED 395  
 DB 600 LFSIWNRRWLRTISILFLNKQDLAEKVLGAGSKIEDYFPEYANTVTPEDATPDAGED 659  
 QY 396 PKVTRAKFFIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVNDCRDI IQRMHLKQY 455  
 DB 660 PKVTRAKFFIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVNDCRDI IQRMHLKQY 719  
 QY 456 ELL 458  
 DB 720 ELL 722

RESULT 14  
 Q5JWF2\_HUMAN PRELIMINARY; PRT; 736 AA.  
 AC Q5JWF2;  
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE GNAS complex locus.  
 GN Names=gnas; ORFNames=RP4-543J19.6-001;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Laird G.;  
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Lloyd D.;  
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Moore M.;  
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL121917; CAI42566.1; -; Genomic DNA.  
 DR EMBL; AL132655; CAI43073.1; -; Genomic DNA.  
 DR EMBL; AL109840; CAI42932.1; -; Genomic DNA.  
 DR EMBL; AL109840; CAI42566.1; JOINED; Genomic DNA.

DR EMBL; AL132655; CAI42566.1; JOINED; Genomic DNA.  
 DR EMBL; AL109840; CAI43073.1; JOINED; Genomic DNA.  
 DR EMBL; AL121917; CAI43073.1; JOINED; Genomic DNA.  
 DR EMBL; AL132655; CAI42932.1; JOINED; Genomic DNA.  
 DR EMBL; AL121917; CAI42932.1; JOINED; Genomic DNA.  
 DR SMR; Q5JWF2; 379-733.  
 DR GO; GO:0005525; F:GTP binding; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPRO11025; GproteinA insert.  
 DR InterPro; IPRO101019; Gprotein\_alpha.  
 DR InterPro; IPRO00367; Gprotein\_alpha\_s.  
 DR InterPro; IPRO02885; PPR.  
 DR Pfam; PF00503; G\_alpha; 1.  
 DR PRINTS; PR00318; GPROTEINA.  
 DR PRINTS; PR00443; GPROTEINAS.  
 DR ProDom; PD000281; Gprotein\_alpha; 1.  
 DR SMART; SM00275; G\_alpha; 1.  
 DR TIGRFAMs; TIGR00756; PPR; 1.  
 SQ SEQUENCE 736 AA; 80219 MW; BD2B6D9A8D00859F CRC64;

Query Match 65.0%; Score 1559; DB 2; Length 736;  
 Best Local Similarity 68.6%; Pred. No. 1.7e-84;  
 Matches 300; Conservative 49; Mismatches 68; Indels 20; Gaps 4;

QY 42 RAAARDARTLLPR---GGEGSPACARPKADKP-KEKRQTEQLSAEEREAAKEREAVKE 97  
 DB 300 RKQPNLLRNFLVQAFGCGFGRSESPQKASRLKVKVPLAEKRQMRQKAELEKRAQKR 359  
 QY 98 A-RKVSIGIDRLMDQKEDLQOQTHRLLLLGAGSGSKSTIVKQRIILHVGFNPE----- 150  
 DB 360 AEKRSKLIDKQLQDEKMGVCTHRLLLLGAGSGSKSTIVKQRIILHVGFNFGEGEEDP 419  
 QY 151 -----EKKQILDIRKNVKDAIVTVSAMSTIIPVPLANPENQFRSDYIKSIAP 201  
 DB 420 QAARSNSDGEKATKVQDIKNLKEAETIVAAASNLPVVELANPENQFRVDYILSVN 479  
 QY 202 TDFEYSOEFFDHVKLWDDEGVKACFRSNEYQLIDCAQYFLERIDSVLSVDYTTDQDL 261  
 DB 480 PDFDFPEFYEYHAKALWEDEGVACFRSNEYQLIDCAQYFLERIDSVLSVDYTTDQDL 539  
 QY 262 LRCRVLTSIGIFETRFQVDKYNFHFVDGGQDERRKWIQCFNDVTAIIYVAACSSYNMVI 321  
 DB 540 LRCRVLTSIGIFETRFQVDKYNFHFVDGGQDERRKWIQCFNDVTAIIYVAACSSYNMVI 599  
 QY 322 REDNTNRLRESLDLFSIWNRRWLRTISILFLNKQDLAEKVLGAGSKIEDYFPEYAN 381  
 DB 600 REDNTNRLRESLDLFSIWNRRWLRTISILFLNKQDLAEKVLGAGSKIEDYFPEYAN 659  
 QY 382 YTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFN 441  
 DB 660 YTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFN 719  
 QY 442 DCRDIIQRMHLKQYELL 458  
 DB 720 DCRDIIQRMHLKQYELL 736

RESULT 15  
 Q5J1N8\_MOUSE PRELIMINARY; PRT; 756 AA.  
 ID Q5J1N8\_MOUSE PRELIMINARY;  
 AC Q5J1N8;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE G-protein XLalphas.  
 GN Name=Gnas;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

Search completed: March 2, 2006, 19:37:28  
Job time : 288.601 secs

**This Page Blank (uspto)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:30:47 ; Search time 35.591 Seconds  
(without alignments)  
1063.905 Million cell updates/sec

Title: US-10-618-320A-1  
Perfect score: 2400  
Sequence: 1 MGLCYSLRPLLCGPGDDPC.....VFNDRCRDIQRMHLKQVELL 458

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1819	75.8	388	2	US-09-949-016-9025
2	1559	65.0	781	2	US-09-949-016-9170
3	1529.5	63.7	801	2	US-08-307-896-1
4	1526.5	63.6	926	2	US-10-314-048A-104
5	1525.5	63.6	394	4	PCT-US95-11808-1
6	1525.5	63.6	869	2	US-10-314-048A-100
7	1522	63.4	775	2	US-09-513-838-6
8	1520.5	63.4	394	2	US-09-442-349A-106
9	1520.5	63.4	1181	2	US-09-826-509-587
10	1464.5	61.0	380	2	US-09-949-016-9251
11	808.5	33.7	374	2	US-09-442-349A-4
12	808.5	33.6	374	2	US-09-442-349A-64
13	805.5	33.6	374	2	US-09-442-349A-70
14	803.5	33.5	374	2	US-09-442-349A-66
15	803.5	33.5	374	2	US-09-442-349A-68
16	803.5	33.5	374	2	US-09-442-349A-69
17	802.5	33.4	374	2	US-09-442-349A-65
18	802.5	33.4	374	2	US-09-442-349A-67
19	802.5	33.4	374	2	US-09-442-349A-76
20	800.5	33.4	374	2	US-09-442-349A-72
21	800.5	33.4	374	2	US-09-442-349A-74
22	800.5	33.4	374	2	US-09-442-349A-75
23	800.5	33.4	374	2	US-09-442-349A-82
24	799.5	33.3	374	2	US-09-442-349A-71
25	799.5	33.3	374	2	US-09-442-349A-79
26	799.5	33.3	374	2	US-09-442-349A-85
27	798.5	33.3	374	2	US-09-442-349A-80

28	798.5	33.3	374	2	US-09-442-349A-81	Sequence 81, Appl
29	797.5	33.2	374	2	US-09-442-349A-77	Sequence 77, Appl
30	797.5	33.2	374	2	US-09-442-349A-78	Sequence 78, Appl
31	797.5	33.2	374	2	US-09-442-349A-83	Sequence 83, Appl
32	797.5	33.2	374	2	US-09-442-349A-84	Sequence 84, Appl
33	797.5	33.2	374	2	US-09-442-349A-90	Sequence 90, Appl
34	796.5	33.2	374	2	US-09-442-349A-73	Sequence 92, Appl
35	796.5	33.2	374	2	US-09-442-349A-92	Sequence 93, Appl
36	795.5	33.1	374	2	US-09-442-349A-87	Sequence 87, Appl
37	795.5	33.1	374	2	US-09-442-349A-89	Sequence 89, Appl
38	794.5	33.1	374	2	US-09-442-349A-86	Sequence 86, Appl
39	794.5	33.1	374	2	US-09-442-349A-88	Sequence 88, Appl
40	794.5	33.1	374	2	US-09-442-349A-91	Sequence 91, Appl
41	745.5	31.1	374	2	US-09-442-349A-3	Sequence 3, Appl
42	742.5	30.9	374	2	US-09-442-349A-35	Sequence 35, Appl
43	742.5	30.9	374	2	US-09-442-349A-41	Sequence 41, Appl
44	740.5	30.9	374	2	US-09-442-349A-37	Sequence 37, Appl
45	740.5	30.9	374	2	US-09-442-349A-39	Sequence 39, Appl

## ALIGNMENTS

### RESULT 1

US-09-949-016-9025  
; Sequence 9025, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9025  
; LENGTH: 388  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9025

Query Match	75.8%	Score 1819;	DB 2;	Length 388;
Best Local Similarity	95.0%	Pred. No. 3.6e-168;		
Matches	345;	Conservative	9;	Mismatches 0; Gaps 0;
Qy	96	KEARKVSRGIDRLMRQKRDLOO	THRLLLGAGESGKSTIVKQMRILHVN	GFPEKKQK 155
Db	26	KERREANKKIEKOLQERLAYKATH	RLLLGAGESGKSTIVKQMRILHVN	GFPEKKQK 85
Qy	156	ILDIRKNVKDAIVTVISAMSTIIP	VPPLANPQNRSDYIKSIAPITDP	EYSOEFPDHVK 215
Db	86	ILDIRKNVKDAIVTVISAMSTIIP	VPPLANPQNRSDYIKSIAPITDP	EYSOEFPDHVK 145
Qy	216	KLWDDGKVACFERSNEYQLIDCA	YFLERIDSVLSVDVYPTDQD	LLRCRLVLTSGIFETR 275
Db	146	KLWDDGKVACFERSNEYQLIDCA	YFLERIDSVLSVDVYPTDQD	LLRCRLVLTSGIFETR 205
Qy	276	FQVDKYNFMFDVGGQDERRKWI	QCFNDVTAIYVAACSSYNNVIRE	DNTNRLRESLD 335
Db	206	FQVDKYNFMFDVGGQDERRKWI	QCFNDVTAIYVAACSSYNNVIRE	DNTNRLRESLD 265
Qy	336	LFESIWNRLRTISILFLNKQDM	LAELKVLACKSKIEDYFPEYANT	VPEDATPDAGD 395
Db	266	LFESIWNRLRTISILFLNKQDM	LAELKVLACKSKIEDYFPEYANT	VPEDATPDAGD 325
Qy	396	PKVTRAKFFIRDLFLRISTATG	DGKHCYCPHFTCAVDTENIRRV	FNDRCRDIQRMHLKQY 455

```
Db 326 PKVTRAKFFIRDLFLRISTATGDKGHCYPHFTCAVDTENIRRVFNCDDIIORMHLKQY 385
Qy 456 ELL 458
Db 386 ELL 388

RESULT 2
US-09-949-016-9170
; Sequence 9170, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9170
; LENGTH: 781
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9170

Query Match 65.08; Score 1559; DB 2; Length 781;
Best Local Similarity 65.6%; Pred. No. 2e-142;
Matches 300; Conservative 49; Mismatches 68; Indels 20; Gaps 4;

Qy 42 RAAARDTARTLLPR--GSGSPACARPKADKP-KEKQRTTEOLSAEEREAEREAVKE 97
Db 345 RKPQNRLLRFLVQAFGCGFRSESPQKASRSLSKVKKVPLAEKRMQRKEALEKRAQKR 404
Qy 98 A-RKVSIGIRMLRDQKDLQOTHRLLLLGAGSGSKSTIVKQMRILHVNCFNPE----- 150
Db 405 AEKRSKLIDKQDEKQGYMCTHRLLLGAGSGSKSTIVKQMRILHVNCFNPEGEEDP 464
Qy 151 -----EKKQKILDIRKNVKDAIIVTSAMSTIIPVPLANPENQFRSDYIKSIAP 201
Db 465 QAARNSDGBKATKVQDIKNLKEAETIIVAAMSNLVPVVELANPENQFRVDYILSVNV 524
Qy 202 TDFEYSOEFFDQVKKLWDDGKVCACFERSNEYQLIDCAQYFLERIDSVLSVDYTPDQDL 261
Db 525 PFDPPPEFYEHAKEALWEDGKVCACFERSNEYQLIDCAQYFLDKIDVIKQADIVPSDQDL 584
Qy 262 LRCRLVLTSGIFETRFQVDKYNFHMFDVGGQDERRRKKWIOCFNDVTATIIYVAACSSYNMVI 321
Db 585 LRCRLVLTSGIFETRFQVDKYNFHMFDVGGQDERRRKKWIOCFNDVTATIIYFVASSYNMVI 644
Qy 322 REDNNTNRLRESLDLPESINNNRWLRTISILFLNKQDMLAEKVLGAGSKIEDYFPEYAN 381
Db 645 REDNQTNRLOEALNLFKSIWNNRWLRTISVILFLNKQDMLAEKVLGAGSKIEDYFPEFAR 704
Qy 382 YTVPEPATPDAGSDPKVTRAKFFIRDLFLRISTATGDKGHCYPHFTCAVDTENIRRVFN 441
Db 705 YTTPEPATPDGSDPKVTRAKFFIRDFLRISTAGDGRHCYPHFTCAVDTENIRRVFN 764
Qy 442 DCRDIIQRMHLKQYELL 458
Db 765 DCRDIIQRMHLRQYELL 781

RESULT 3
US-08-307-896-1
; Sequence 1, Application US/08307896C
```

```
; Patent No. 6034071
; GENERAL INFORMATION:
; APPLICANT: Iyengar, Srinivas Ravi
; TITLE OF INVENTION: MUTANT ACTIVATED GSPALPHA AND ADENYLYL
; FILE REFERENCE: 29770
; CURRENT APPLICATION NUMBER: US/08/307,896C
; CURRENT FILING DATE: 1994-09-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-307-896-1

Query Match 63.7%; Score 1529.5; DB 2; Length 380;
Best Local Similarity 76.2%; Pred. No. 5e-140;
Matches 281; Conservative 45; Mismatches 42; Indels 1; Gaps 1;

Qy 91 EREAVKEARKVSRGIDRMLRDQKDLQOTHRLLLLGAGSGSKSTIVKQMRILHVNCFNPE 150
Db 12 QRNEEKAQREANKIEKQLQDKQVYRATHRLLLGAGSGSKSTIVKQMRILHVNCFNGD 71
Qy 151 -EKKQKILDIRKNVKDAIIVTSAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSOE 209
Db 72 SEKATKVQDIKNLKEAETIIVAAMSNLVPVVELANPENQFRVDYILSVNVDFDPPPE 131
Qy 210 FFDHVKKLWDDGKVCACFERSNEYQLIDCAQYFLERIDSVLSVDYTPDQDLRLCRVLTS 269
Db 132 FYEHAKALWEDGKVCACFERSNEYQLIDCAQYFLDKIDVIKQADYVPSDQDLRLCRVLTS 191
Qy 270 GIFETRFQVDKYNFHMFDVGGQDERRRKKWIOCFNDVTATIIYVAACSSYNMVIREDNNTNR 329
Db 192 GIFETRFQVDKYNFHMFDVGGQDERRRKKWIOCFNDVTATIIYFVASSYNMVIREDNQTNR 251
Qy 330 LRESLDLPESINNNRWLRTISILFLNKQDMLAEKVLGAGSKIEDYFPEYANYTVPEPAT 389
Db 252 LQEALNLFKSIWNNRWLRTISVILFLNKQDMLAEKVLGAGSKIEDYFPEFARYTTPEPAT 311
Qy 390 PDAGSDPKVTRAKFFIRDLFLRISTATGDKGHCYPHFTCAVDTENIRRVFNCDDIIQ 449
Db 312 PEPGSDPKVTRAKFFIRDFLRISTAGDGRHCYPHFTCAVDTENIRRVFNCDDIIQ 371
Qy 450 MHLKQYELL 458
Db 372 MHLRQYELL 380

RESULT 4
US-10-314-048A-104
; Sequence 104, Application US/10314048A
; Patent No. 6902902
; GENERAL INFORMATION:
; APPLICANT: Unett, David J.
; APPLICANT: Chen, Ruoping
; APPLICANT: Richman, Jeremy
; APPLICANT: Connolly, Daniel
; APPLICANT: Dang, Huong T.
; APPLICANT: Choi, Bryan
; APPLICANT: Leonard, James
; APPLICANT: Hakak, Yaron
; APPLICANT: Liaw, Chen
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Lerner, Michael
; TITLE OF INVENTION: Human G Protein-Coupled Receptors and Modulators Thereof
; FILE REFERENCE: 22.US6.CIP
; CURRENT APPLICATION NUMBER: US/10/314,048A
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/096,511
```

```

; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 08/995,543
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/399,917
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/404,761
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 60/410,747
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 104
; LENGTH: 926
; TYPE: PRT
; ORGANISM: Homo sapiens and Rat
; US-10-314-048A-104

Query Match      63.6%; Score 1526.5; DB 2; Length 926;
Best Local Similarity 63.5%; Pred. No. 3.8e-139;
Matches 298; Conservative 51; Mismatches 67; Indels 53; Gaps 5;

QY 24 EPPVEDAQP-----APAPALAPVRAAARDTARTLL-----PRGGEGSPACA 64
DB 477 KPPKEDSHDLPCTEGTEGKIVPSYDSATFPALSAEFHHTGLVDPSVSLGCRSNGC- 535
QY 65 RPKADPKKQKQTEQLSABEREAAKREAVKRGVSGIDRMLRDKQKDLQOQTHRLLL 124
DB 536 ---LGNSKTEDQRNEE-----KAQREANKKIEKQLQKDKQVYRATHRLLL 577
QY 125 LGAGESGKSTIVKQMRILHVNNGNPE-----EKKQKILDIRKVNKDAIVT 169
DB 578 LGAGESGKSTIVKQMRILHVNNGNPE-----EKKQKILDIRKVNKDAIVT 169
QY 170 IVSAMSTIIPVPLANPENOFSDYIKSIAPITDFEYSQBFDFHVKKLWDEGVKACFER 229
DB 638 IVAMSNLVPVELANPENOFSDYIILSVNVPNDFPPPEFYEHAKALWEDEGVACFER 697
QY 230 SNEYQLIDCAQYFLERIDSVLDVYPTDQDLRCRLVLTSGIFETRFQVDKVNFMFDVG 289
DB 698 SNEYQLIDCAQYFLERIDSVLDVYPTDQDLRCRLVLTSGIFETRFQVDKVNFMFDVG 757
QY 290 GQDERKWTQCNVDVTALIVVAACSSYNMVIREDNNTNRLRESLDLFPESIWNNRMLRTI 349
DB 758 GQDERKWTQCNVDVTALIVVAACSSYNMVIREDNNTNRLRESLDLFPESIWNNRMLRTI 817
QY 350 SIILFLNKQDLAEKVLAKGSKIEDYYPEYANTVTPEDATPDAGEDPKVTRAKFFIRDLF 409
DB 818 SVILFLNKQDLAEKVLAKGSKIEDYYPEYANTVTPEDATPDAGEDPKVTRAKFFIRDLF 877
QY 410 LRISTATGDKGHYCPHFTCAVDTENIRRVFNDCRDIIRMHKQYELL 458
DB 878 LRISTATGDKGHYCPHFTCAVDTENIRRVFNDCRDIIRMHKQYELL 926

RESULT 5
PCT-US95-11808-1
; Sequence 1, Application PC/TUS9511808
; GENERAL INFORMATION:
; APPLICANT: Iyengar, Srinivas Ravi V.
; TITLE OF INVENTION: MUTANT ACTIVATED GsALPHA AND ADENYLYL
; TITLE OF INVENTION: ADENYLYL
; CYCLOTYPE OF INVENTION: CYCLASE 2 FOR USE AS THERAPEUTIC AGENTS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue and
; ADDRESSEE: Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11808
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/307,896
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S.
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: 29970 165/28755
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2500
; TELEFAX: (212) 765-2519
; TELEX: 650 6111063
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: GsALPHA
; PCT-US95-11808-1

Query Match      63.6%; Score 1525.5; DB 4; Length 394;
Best Local Similarity 73.6%; Pred. No. 1.3e-139;
Matches 282; Conservative 44; Mismatches 42; Indels 15; Gaps 1;

QY 91 EREAVAEARKVSGIDRMLRDKQKDLQOQTHRLLLGAGESGKSTIVKQMRILHVNNGNPE 150
DB 12 QRNEEKAQREANKKIEKQLQKDKQVYRATHRLLLGAGESGKSTIVKQMRILHVNNGNPE 71
QY 151 -----EKKQKILDIRKVNKDAIVTIVSAMSTIIPVPLANPENQFRSDYI 195
DB 72 GGEEDPQAARSNSDGEKATKVQDIKNLKEAIEITIVAAMSNLVPVELANPENQFRVDYI 131
QY 196 KSTAPITDFEYSQBFDFHVKKLWDEGVKACFERSENEYQLIDCAQYFLERIDSVLDVYT 255
DB 132 LSVNVPNDFPPPEFYEHAKALWEDEGVACFERSENEYQLIDCAQYFLERIDSVLDVYT 191
QY 256 PTQDQLRCRLVLTSGIFETRFQVDKVNFMFDVGQDERKWTQCNVDVTALIVVAACS 315
DB 192 PSDQDLRCRLVLTSGIFETRFQVDKVNFMFDVGQDERKWTQCNVDVTALIVVAACS 251
QY 316 SYNVMVIREDNNTNRLRESLDLFPESIWNNRMLRTISILFLNKQDLAEKVLAKGSKIEDY 375
DB 252 SYNVMVIREDNNTNRLRESLDLFPESIWNNRMLRTISILFLNKQDLAEKVLAKGSKIEDY 311
QY 376 FPEYANTVTPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDKGHYCPHFTCAVDTEN 435
DB 312 FPEYANTVTPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDKGHYCPHFTCAVDTEN 371
QY 436 IRRVFNDCRDIIRMHKQYELL 458
DB 372 IRRVFNDCRDIIRMHKQYELL 394

RESULT 6
US-10-314-048A-100
; Sequence 100, Application US/10314048A
; Patent No. 6902902
; GENERAL INFORMATION:
; APPLICANT: Unett, David J.
; APPLICANT: Chen, Ruoping
; APPLICANT: Richman, Jeremy
; APPLICANT: Connolly, Daniel
; APPLICANT: Dang, Huong T.
; APPLICANT: Choi, Bryan

```

; APPLICANT: Leonard, James  
 ; APPLICANT: Hakak, Yaron  
 ; APPLICANT: Liaw, Chen  
 ; APPLICANT: Lowitz, Kevin P.  
 ; APPLICANT: Behan, Dominic P.  
 ; APPLICANT: Chalmers, Derek T.  
 ; APPLICANT: Lerner, Michael  
 ; TITLE OF INVENTION: Human G Protein-Coupled Receptors and Modulators Thereof  
 ; TITLE OF INVENTION: for the Treatment of Metabolic-Related Disorders  
 ; FILE REFERENCE: 22 US6 CIP  
 ; CURRENT APPLICATION NUMBER: US/10/314,048A  
 ; CURRENT FILING DATE: 2002-12-06  
 ; PRIOR APPLICATION NUMBER: 10/096,511  
 ; PRIOR FILING DATE: 2002-03-12  
 ; PRIOR APPLICATION NUMBER: 09/995,543  
 ; PRIOR FILING DATE: 2001-11-27  
 ; PRIOR APPLICATION NUMBER: 60/399,917  
 ; PRIOR FILING DATE: 2002-07-29  
 ; PRIOR APPLICATION NUMBER: 60/404,761  
 ; PRIOR FILING DATE: 2002-08-19  
 ; PRIOR APPLICATION NUMBER: 60/410,747  
 ; PRIOR FILING DATE: 2002-09-13  
 ; NUMBER OF SEQ ID NOS: 161  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 100  
 ; LENGTH: 869  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens and Rat  
 ; US-10-314-048A-100

Query Match 63.6%; Score 1525.5; DB 2; Length 869;  
 Best Local Similarity 60.5%; Pred. No. 4.4e-139;  
 Matches 302; Conservative 52; Mismatches 70; Indels 75; Gaps 7;

Qy	8	RPLFGGDDPCAASPPVED-----	-----AQPAPAL 38
Db	398	RPL-----PSPKQPPAVDFRIPQIAETSEFLEQQLTSDTIIMSDSYLRPAASRL 449	
Qy	39	APVRAARDTA---RLLPRGEGSPACARPKADPKKQRTQLSAEREAAKERE 94	
Db	450	ESAISAEEHTGLVDPSSVSLGCRSMGC-----LGNSKTDQNEE----- 491	
Qy	95	VKEARKVSRGIDRLMRDQKRDLOQTHRLLLGAGESGKSTIVKQMRILHVGPNPE----- 150	
Db	492	-KAQREANKIKQLQKDKQVYRATHRLLLGAGESGKSTIVKQMRILHVGNGEGGEE 550	
Qy	151	-----EKKQKILDIRKNVDAIVTVISAMSTIIPVPLANPENQFRSDYIKSIA 199	
Db	551	DPQAARSNDGKATKQDQIKNNLKEAETIVAAMSNLVPVELANPENQFRVDYILSYM 610	
Qy	200	PITDPEYSQEFFDHVKLWDEGVKACFERSNEYQLIDCAQYFLERIDSVSLVDTPTDQ 259	
Db	611	NVPNDFDPPEFYEHAKEALWEDEGRACYSERSNEYQLIDCAQYFLDKIDVIKQADYVPSDQ 670	
Qy	260	DLRCRVLTSGIFETRFQVDKVNFMFVGGQDERRKWIQCNDVTAIIYVAACSSYNN 319	
Db	671	DLRCRVLTSGIFETRFQVDKVNFMFVGGQDERRKWIQCNDVTAIIYVAACSSYNN 730	
Qy	320	VIREDNNTLRSLDLFESINWNRWLTISIILFNKQDLAEKVLGKSKIEDYFPEY 379	
Db	731	VIREDNNTLRSLDLFESINWNRWLTISIILFNKQDLAEKVLGKSKIEDYFPEY 790	
Qy	380	ANYTVPEDATPDGADPKVTRAKFFRLDLRLSTATGDKHYCPHFTCAVDTENIRRV 439	
Db	791	ARYTTPEDATPEGEDPRVTRAKYFIRDEFRLSTATSGDGRHYCPHFTCAVDTENIRRV 850	
Qy	440	FNDCRDIIQRMHLKQYELL 458	
Db	851	FNDCRDIIQRMHLKQYELL 869	

; Sequence 6, Application US/09513838  
 ; Patent No. 6420563  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beeley, Nigel R  
 ; APPLICANT: Behan, Dominic P  
 ; APPLICANT: Chalmers, Derek T  
 ; APPLICANT: Menzaghi, Frederique  
 ; APPLICANT: Strah-Pleynt, Sonja  
 ; TITLE OF INVENTION: Small Molecule Modulators Of G Protein-Coupled Receptor  
 ; TITLE OF INVENTION: Six  
 ; FILE REFERENCE: AREN0058  
 ; CURRENT APPLICATION NUMBER: US/09/513,838  
 ; CURRENT FILING DATE: 2000-02-25  
 ; EARLIER APPLICATION NUMBER: 09/364,425  
 ; EARLIER FILING DATE: 1999-07-30  
 ; EARLIER APPLICATION NUMBER: 60/094,879  
 ; EARLIER FILING DATE: 1998-07-31  
 ; EARLIER APPLICATION NUMBER: 60/106,300  
 ; EARLIER FILING DATE: 1998-10-30  
 ; EARLIER APPLICATION NUMBER: 60/110,906  
 ; EARLIER FILING DATE: 1998-12-04  
 ; EARLIER APPLICATION NUMBER: 60/121,851  
 ; EARLIER FILING DATE: 1999-02-26  
 ; EARLIER APPLICATION NUMBER: 60/173,850  
 ; EARLIER FILING DATE: 1999-12-30  
 ; EARLIER APPLICATION NUMBER: 60/174,428  
 ; EARLIER FILING DATE: 2000-01-04  
 ; EARLIER APPLICATION NUMBER: 09/364,425  
 ; EARLIER FILING DATE: 1999-07-30  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 775  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-513-838-6

Query Match	63.4%;	Score 1522;	DB 2;	Length 775;
Best Local Similarity	67.4%;	Pred. No. 8e-139;		
Matches	289;	Conservative 49;	Mismatches 57;	Indels 34; Gaps 3;
Qy	45	ARDTARTLLPRGEGSPACARPKADPKKQRTQLSAEREAAKEREAVKARKVSG 104		
Db	366	SRSPEVRVPSLGRSMGC-----LGNSKTDQNEE-----KAQREANKK 406		
Qy	105	IDRLMRDQKRDLOQTHRLLLGAGESGKSTIVKQMRILHVGPNPE----- 150		
Db	407	IEKQLQKDKQVYRATHRLLLGAGESGKSTIVKQMRILHVGNGEGGEEPPQAARSND 466		
Qy	151	-EKKQKILDIRKNVDAIVTVISAMSTIIPVPLANPENQFRSDYIKSITAPITDPEYSOE 209		
Db	467	GKATKQDQIKNNLKEAETIVAAMSNLVPVELANPENQFRVDYILSVNVPNDFPPE 526		
Qy	210	FPDHVKLWDEGVKACFERSNEYQLIDCAQYFLERIDSVSLVDTPTDQDLRCRVLT 269		
Db	527	FYEHAKEALWEDEGRACYSERSNEYQLIDCAQYFLDKIDVIKQADYVPSDQDLRCRVLT 586		
Qy	270	GIFETRFQVDKVNFMFVGGQDERRKWIQCNDVTAIIYVAACSSYNNVIREDNNTNR 329		
Db	587	GIFETRFQVDKVNFMFVGGQDERRKWIQCNDVTAIIYVAACSSYNNVIREDNNTNR 646		
Qy	330	LRSLDLFESINWNRWLTISIILFNKQDLAEKVLGKSKIEDYFPEYANYTVPEDAT 389		
Db	647	LRSLDLFESINWNRWLTISIILFNKQDLAEKVLGKSKIEDYFPEYANYTVPEDAT 706		
Qy	390	PDAGDPKVTAKFFIRDLRLSTATGDKHYCPHFTCAVDTENIRRVFNDCRDIIQR 449		
Db	707	PDAGDPKVTAKFFIRDLRLSTATGDKHYCPHFTCAVDTENIRRVFNDCRDIIQR 766		
Qy	450	MHLKQYELL 458		
Db	767	MHLKQYELL 775		

[illegible]

```
QY 217 LWDEGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYPTDQDLRCRLVLTSGIPETRF 276
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 139 LWDEGVKACFERSNEYQLIDCAQYFLDKDVIKQADYVPSDQDLRCRLVLTSGIPETRF 198
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 277 QVDKVNFMFDVGGQDERRKWTQCFNDVTAIIVAAACSSYNMVRREDNNNRLRESLDL 336
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 199 QVDKVNFMFDVGGQDERRKWTQCFNDVTAIIVVASSYNMVRREDNNQTLQALNL 258
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 337 FESIMNNRWLRTISILFLNKQDLAEKVLAKGSKIEDYPEYANVTVPEDATPDGEDP 396
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 259 FKSIMNNRWLRTISVILFLNKQDLAEKVLAKGSKIEDYPEYFARYTTPEDATPEGEDP 318
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 397 KVTAKFFIRDLRLISTATGDKGHYCYPHFTCAVDTENIRRVFNDRCRDIIOQMHLKQYE 456
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 319 RVTRAKVFIREFLRISTASGDRGHYCYPHFTCAVDTENIRRVFNDRCRDIIOQMHLRQYE 378
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 457 LL 458
||
Db 379 LL 380
||

RESULT 11
US-09-442-349A-4
; Sequence 4, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
; OTHER INFORMATION: Chimera
US-09-442-349A-4

Query Match 33.7%; Score 808.5; DB 2; Length 374;
Best Local Similarity 44.0%; Pred. No. 6.3e-70;
Matches 166; Conservative 71; Mismatches 111; Indels 29; Gaps 6;

QY 93 EAVKEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGESGKSTIVKQMRIHVNFGNPEEK 152
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 16 EDEKAAARVDQEIINRILLEKQKQDRGELKLLLLGPGESGKSTIKQMRIHAGYSEER 75
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 153 KQKILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAP--ITDFEYSQEF 210
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 76 KGFRPLVYQNIIVFSMRAMIEAMERL--QIPFSPESKHHASLVMSQDPYKVTTFE--KRY 131
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 211 FDHVKKLWDDEGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYPTDQDLRCRLVLTSG 270
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 132 AAAMQWLWDAGIRACYERRERFHLDSAVYVLSHLERITEEGYVPTAQDVLRSRPTTG 191
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 271 IFETRFQVDKVNFMFDVGGQDERRKWTQCFNDVTAIIVAAACSSYNMVRREDNNNRL 330
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 192 INEYCFSVQKTNLRIVDVGQKSERKKWICFENVIALIYVLAISLSEYDQCLENNQENRM 251
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 331 RESLDLFEFESIMNNRWLRTISILFLNKQDLAEKVLAKGSKIEDYPEYANVTVPEDATP 390
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 252 KESLALFGTILELPWFKSTSVILFLNKTDLLEKI--PTSHLATYFPSP----- 298
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 391 DAGEDPKVTRAKFFIRDLRLIST-----ATGDGKHVCYPHFTCAVDTENIRRVFN 441
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 299 -QGPKQDAEAAKRFILDMYTRMYTCVDGPEGSASGDGRHVCYPHFTCAVDTENIRRVFN 357
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 442 DCRDIIOQMHLKQYELL 458
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 358 DCRDIIOQMHLRQYELL 374
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 13
US-09-442-349A-70
; Sequence 70, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
Db 358 DCRDIIOQMHLRQYELL 374
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 12
US-09-442-349A-64
; Sequence 64, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
; OTHER INFORMATION: Chimera
US-09-442-349A-64

Query Match 33.6%; Score 805.5; DB 2; Length 374;
Best Local Similarity 43.8%; Pred. No. 1.2e-69;
Matches 165; Conservative 72; Mismatches 111; Indels 29; Gaps 6;

QY 93 EAVKEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGESGKSTIVKQMRIHVNFGNPEEK 152
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 16 EDEKAAARVDQEIINRILLEKQKQDRGELKLLLLGPGESGKSTIKQMRIHAGYSEER 75
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 153 KQKILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAP--ITDFEYSQEF 210
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 76 KGFRPLVYQNIIVFSMRAMIEAMERL--QIPFSPESKHHASLVMSQDPYKVTTFE--KRY 131
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 211 FDHVKKLWDDEGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYPTDQDLRCRLVLTSG 270
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 132 AAAMQWLWDAGIRACYERRERFHLDSAVYVLSHLERITEEGYVPTAQDVLRSRPTTG 191
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 271 IFETRFQVDKVNFMFDVGGQDERRKWTQCFNDVTAIIVAAACSSYNMVRREDNNNRL 330
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 192 INEYCFSVQKTNLRIVDVGQKSERKKWICFENVIALIYVLAISLSEYDQCLENNQENRM 251
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 331 RESLDLFEFESIMNNRWLRTISILFLNKQDLAEKVLAKGSKIEDYPEYANVTVPEDATP 390
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 252 KESLALFGTILELPWFKSTSVILFLNKTDLLEKI--PTSHLATYFPSP----- 298
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 391 DAGEDPKVTRAKFFIRDLRLIST-----ATGDGKHVCYPHFTCAVDTENIRRVFN 441
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 299 -QGPKQDAEAAKRFILDMYTRMYTCVDGPEGSASGDGRHVCYPHFTCAVDTENIRRVFN 357
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 442 DCRDIIOQMHLKQYELL 458
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 358 DCRDIIOQMHLRQYELL 374
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 13
US-09-442-349A-70
; Sequence 70, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: G protein  
; OTHER INFORMATION: Chimera  
US-09-442-349A-70

Query Match 33.6%; Score 805.5; DB 2; Length 374;  
Best Local Similarity 43.8%; Pred. No. 1.2e-69;  
Matches 165; Conservative 72; Mismatches 111; Indels 29; Gaps 6;  
QY 93 EAVKEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGESGKSTIVKQRIILHVGFPNPEEK 152  
DB 16 EDEKAAARVDQENIRLLEQKQDQRGELKLLLLGPGESGKSTFIKQRIIHGAGYSEER 75  
QY 153 KQKILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAP--ITDFEYSQEF 210  
DB 76 KGRPLVYQNI FVSMRAMIEMERL--QIPFSPESKHASLVMSQDPYKVTTFE--KRY 131  
QY 211 FDHVKKLWDDGKACFERNEVQLIDCAQFYLERIDSVLSVDYPTDQDLLRCRVLTSG 270  
DB 132 AAAMQWLWDAGIRACVERREPHLLDSAVYYLSHLERITEEGVYPTAQDVLRSRMTTG 191  
QY 271 IFETRFQVDKVNPFHMFVGGQDERRRKKWIOCFNDVTAIYVAAACSSNMVIREDDNNTNRL 330  
DB 192 INEVCFSVQKTNLRIVDVGQSKERKKWIHCFENVIALIYLSLSEYDQCLEENQENRM 251  
QY 331 RESLDLFESINWNRWLTISIILFLNKQDMLAEKVLGAGSKI EDYFPEYANYTVPEDATP 390  
DB 252 KESLALFGTILELPWFKSTSVILFNKTDILEEKI--PTSHLATYFPSP----- 298  
QY 391 DAGEDPKVTRAKFFIRDLFLRIST-----ATDGGKHVCYPHFTCAVDTENIRRVFN 441  
DB 299 -QQPKQDAEAAKRFILDMYTRMTGCVDPGEGSASGDRHRCYPHFTCAVDTENIRRVFN 357  
QY 442 DCRDIIOHMLKQYELL 458  
DB 358 DCRDIIOHMLKQYELL 374

## RESULT 14

US-09-442-349A-66  
; Sequence 66; Application US/09442349A  
; Patent No. 6462178  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Yung H  
; TITLE OF INVENTION: G Protein  
; FILE REFERENCE: M99/0101/US  
; CURRENT APPLICATION NUMBER: US/09/442,349A  
; CURRENT FILING DATE: 1999-11-17  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 66  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: G protein  
; OTHER INFORMATION: Chimera  
US-09-442-349A-66

Query Match 33.5%; Score 803.5; DB 2; Length 374;  
Best Local Similarity 43.8%; Pred. No. 1.9e-69;  
Matches 165; Conservative 72; Mismatches 111; Indels 29; Gaps 6;  
QY 93 EAVKEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGESGKSTIVKQRIILHVGFPNPEEK 152  
DB 16 EDEKAAARVDQENIRLLEQKQDQRGELKLLLLGPGESGKSTFIKQRIIHGAGYSEER 75  
QY 153 KQKILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAP--ITDFEYSQEF 210  
DB 76 KGRPLVYQNI FVSMRAMIEMERL--QIPFSPESKHASLVMSQDPYKVTTFE--KRY 131  
QY 211 FDHVKKLWDDGKACFERNEVQLIDCAQFYLERIDSVLSVDYPTDQDLLRCRVLTSG 270

DB 132 AAAMQWLWDAGIRACVERREPHLLDSAVYYLSHLERITEEGVYPTAQDVLRSRMTTG 191  
QY 271 IFETRFQVDKVNPFHMFVGGQDERRRKKWIOCFNDVTAIYVAAACSSNMVIREDDNNTNRL 330  
DB 192 INEVCFSVQKTNLRIVDVGQSKERKKWIHCFENVIALIYLSLSEYDQCLEENQENRM 251  
QY 331 RESLDLFESINWNRWLTISIILFLNKQDMLAEKVLGAGSKI EDYFPEYANYTVPEDATP 390  
DB 252 KESLALFGTILELPWFKSTSVILFNKTDILEEKI--PTSHLATYFPSP----- 298  
QY 391 DAGEDPKVTRAKFFIRDLFLRIST-----ATDGGKHVCYPHFTCAVDTENIRRVFN 441  
DB 299 -QQPKQDAEAAKRFILDMYTRMTGCVDPGEGSASGDRHRCYPHFTCAVDTENIRRVFN 357  
QY 442 DCRDIIOHMLKQYELL 458  
DB 358 DCRDIIOHMLKQYELL 374  
RESULT 15  
US-09-442-349A-68  
; Sequence 68; Application US/09442349A  
; Patent No. 6462178  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Yung H  
; TITLE OF INVENTION: G Protein  
; FILE REFERENCE: M99/0101/US  
; CURRENT APPLICATION NUMBER: US/09/442,349A  
; CURRENT FILING DATE: 1999-11-17  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 68  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: G protein  
; OTHER INFORMATION: Chimera  
US-09-442-349A-68

Query Match 33.5%; Score 803.5; DB 2; Length 374;  
Best Local Similarity 43.8%; Pred. No. 1.9e-69;  
Matches 165; Conservative 71; Mismatches 112; Indels 29; Gaps 6;  
QY 93 EAVKEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGESGKSTIVKQRIILHVGFPNPEEK 152  
DB 16 EDEKAAARVDQENIRLLEQKQDQRGELKLLLLGPGESGKSTFIKQRIIHGAGYSEER 75  
QY 153 KQKILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAP--ITDFEYSQEF 210  
DB 76 KGRPLVYQNI FVSMRAMIEMERL--QIPFSPESKHASLVMSQDPYKVTTFE--KRY 131  
QY 211 FDHVKKLWDDGKACFERNEVQLIDCAQFYLERIDSVLSVDYPTDQDLLRCRVLTSG 270  
DB 132 AAAMQWLWDAGIRACVERREPHLLDSAVYYLSHLERITEEGVYPTAQDVLRSRMTTG 191  
QY 271 IFETRFQVDKVNPFHMFVGGQDERRRKKWIOCFNDVTAIYVAAACSSNMVIREDDNNTNRL 330  
DB 192 INEVCFSVQKTNLRIVDVGQSKERKKWIHCFENVIALIYLSLSEYDQCLEENQENRM 251  
QY 331 RESLDLFESINWNRWLTISIILFLNKQDMLAEKVLGAGSKI EDYFPEYANYTVPEDATP 390  
DB 252 KESLALFGTILELPWFKSTSVILFNKTDILEEKI--PTSHLATYFPSP----- 298  
QY 391 DAGEDPKVTRAKFFIRDLFLRIST-----ATDGGKHVCYPHFTCAVDTENIRRVFN 441  
DB 299 -QQPKQDAEAAKRFILDMYTRMTGCVDPGEGSASGDRHRCYPHFTCAVDTENIRRVFN 357  
QY 442 DCRDIIOHMLKQYELL 458  
DB 358 DCRDIIOHMLKQYELL 374



Search completed: March 2, 2006, 19:31:49  
Job time : 35.591 secs

---

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:32:31 ; Search time 208.863 Seconds  
(without alignments)  
916.227 Million cell updates/sec

Title: US-10-618-320A-1

Perfect score: 2400

Sequence: 1 MGLCYSLRPLLFQGGPDDPC.....VFNDCRDIQRMHLKQYELL 458

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2400	100.0	458	5	US-10-732-923-8011
2	1847	77.0	351	5	US-10-732-923-8014
3	1819	75.8	381	3	US-09-952-680A-18
4	1819	75.8	381	5	US-10-215-982-18
5	1819	75.8	381	5	US-10-732-923-7615
6	1819	75.8	381	5	US-10-732-923-8012
7	1819	75.8	381	5	US-10-736-149-5241
8	1811	75.5	381	5	US-10-732-923-7651
9	1677	69.9	379	5	US-10-732-923-7963
10	1559	65.0	756	3	US-09-963-131-184
11	1559	65.0	756	5	US-10-732-923-7617
12	1559	65.0	909	5	US-10-732-923-8020
13	1548	64.5	715	5	US-10-732-923-7654
14	1548	64.5	846	5	US-10-732-923-7655
15	1541	64.2	846	5	US-10-732-923-7656
16	1540	64.2	379	3	US-09-952-680A-15
17	1540	64.2	379	4	US-10-408-765A-105
18	1540	64.2	379	5	US-10-215-982-15
19	1540	64.2	379	5	US-10-732-923-8015
20	1537	64.0	720	4	US-10-164-163-16
21	1530	63.7	377	5	US-10-732-923-7618
22	1529.5	63.7	380	3	US-09-952-680A-16
23	1529.5	63.7	380	5	US-10-215-982-16
24	1529.5	63.7	380	5	US-10-732-923-8016
25	1527	63.6	379	5	US-10-732-923-7964
26	1526.5	63.6	380	5	US-10-732-923-8055
27	1526.5	63.6	388	5	US-10-732-923-8054

28	1526.5	63.6	926	4	US-10-321-807-104	Sequence 104, App
29	1526.5	63.6	926	4	US-10-321-807-104	Sequence 104, App
30	1526.5	63.6	926	4	US-10-314-048A-104	Sequence 104, App
31	1526.5	63.6	926	5	US-10-897-815-104	Sequence 104, App
32	1526.5	63.6	926	5	US-10-930-662-104	Sequence 104, App
33	1525.5	63.6	394	3	US-09-952-680A-13	Sequence 13, Appl
34	1525.5	63.6	394	3	US-09-963-131-194	Sequence 134, App
35	1525.5	63.6	394	4	US-10-116-275-187	Sequence 187, App
36	1525.5	63.6	394	4	US-10-352-843-18	Sequence 18, Appl
37	1525.5	63.6	394	5	US-10-215-982-13	Sequence 13, Appl
38	1525.5	63.6	394	5	US-10-684-422-194	Sequence 194, App
39	1525.5	63.6	394	5	US-10-732-923-8017	Sequence 8017, App
40	1525.5	63.6	394	5	US-10-732-923-8019	Sequence 8019, App
41	1525.5	63.6	869	4	US-10-321-807-100	Sequence 100, App
42	1525.5	63.6	869	4	US-10-321-807-100	Sequence 100, App
43	1525.5	63.6	869	4	US-10-314-048A-100	Sequence 100, App
44	1525.5	63.6	869	5	US-10-897-815-100	Sequence 100, App
45	1525.5	63.6	869	5	US-10-930-662-100	Sequence 100, App

## ALIGNMENTS

## RESULT 1

US-10-732-923-8011

; Sequence 8011, Application US/10732923

; Publication No. US20050108791A1

; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D

; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

; FILE REFERENCE: 38-15(52796)C

; CURRENT APPLICATION NUMBER: US/10/732,923

; PRIOR FILING DATE: 2003-12-10

; PRIOR APPLICATION NUMBER: 10/310,154

; PRIOR FILING DATE: 2002-12-04

; NUMBER OF SEQ ID NOS: 24149

; SEQ ID NO 8011

; LENGTH: 458

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-732-923-8011

Query Match 100.0%; Score 2400; DB 5; Length 458;  
Best Local Similarity 100.0%; Pred. No. 1.2e-174;  
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGLCYSLRPLLFQGGPDDPCAASEPVEDAQAPALAPVRAAARDTARTLLPRGEGS	60
Db	1	MGLCYSLRPLLFQGGPDDPCAASEPVEDAQAPALAPVRAAARDTARTLLPRGEGS	60
Qy	61	PACARPKADKPKRKORTQOLSAEEREAAKREAVKEARKVSGIDRMLDKRDLQOOTH	120
Db	61	PACARPKADKPKRKORTQOLSAEEREAAKREAVKEARKVSGIDRMLDKRDLQOOTH	120
Qy	121	RLLLGAGSGKSTIVKQMRILHVNGFNPEKKQKILDIRNVKDAIVTVSVAMSTIIPP	180
Db	121	RLLLGAGSGKSTIVKQMRILHVNGFNPEKKQKILDIRNVKDAIVTVSVAMSTIIPP	180
Qy	181	VPLANPENQPSDIYIKSIAPITDFEYSQEPFDHVKKLWDDGEGVACPFERNVQLIDCAQ	240
Db	181	VPLANPENQPSDIYIKSIAPITDFEYSQEPFDHVKKLWDDGEGVACPFERNVQLIDCAQ	240
Qy	241	YFLERIDSVLDVYTPDQDLRCRVLTSGIFETRFQVDKVNFMFVGGQRRERKWIQ	300
Db	241	YFLERIDSVLDVYTPDQDLRCRVLTSGIFETRFQVDKVNFMFVGGQRRERKWIQ	300
Qy	301	CFNDVTAIYVAACSSYNMVIREDNNNRLRESLDLFESINNNRLRTISIIILFLNKQDM	360
Db	301	CFNDVTAIYVAACSSYNMVIREDNNNRLRESLDLFESINNNRLRTISIIILFLNKQDM	360
Qy	361	LAEKVLGSKSIBDYFPEYANYTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK	420
Db	361	LAEKVLGSKSIBDYFPEYANYTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTATGDGK	420

```
QY 421 HVCYPHFTCAVDTENIRRVNDCRDIIQRMHLKQYELL 458
|||||
Db 421 HVCYPHFTCAVDTENIRRVNDCRDIIQRMHLKQYELL 458
|||||
RESULT 2
US-10-732-923-8014
; Sequence 8014, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR FILING DATE: 2003-12-10
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8014
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-732-923-8014
Query Match 77.0%; Score 1847; DB 5; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.4e-132; Indels 0; Gaps 0;
Matches 351; Conservative 0; Mismatches 0;
QY 108 MLRQKRDLOQTHRLLLGAGESGKSTIVQMRLHVNFGNPNPEKKQKILDIRKNVKDAI 167
|||||
Db 1 MLRQKEDLOQTHRLLLGAGESGKSTIVQMRLHVNFGNPNPEKKQKILDIRKNVKDAI 60
|||||
QY 168 VTIIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSQEFDFHVKLWDDGKACF 227
|||||
Db 61 VTIIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSQEFDFHVKLWDDGKACF 120
|||||
QY 228 ERSNEYQLIDCAQYFLERIDSVLVDTPTDQDLRLCRVLTSGIFETRFQVDKVNHEMD 287
|||||
Db 121 ERSNEYQLIDCAQYFLERIDSVLVDTPTDQDLRLCRVLTSGIFETRFQVDKVNHEMD 180
|||||
QY 288 VGGORDERRKWIQCFNDVTAIIVAACSSYNMVIREDNNTNRLRESLDLPESIWNNRWLR 347
|||||
Db 181 VGGORDERRKWIQCFNDVTAIIVAACSSYNMVIREDNNTNRLRESLDLPESIWNNRWLR 240
|||||
QY 348 TISILFLNKQDMLAEKVLGAGSKIEDYFPEYANYTVPEDATPDAGEDPKVTRAKFFIRD 407
|||||
Db 241 TISILFLNKQDMLAEKVLGAGSKIEDYFPEYANYTVPEDATPDAGEDPKVTRAKFFIRD 300
|||||
QY 408 LFURISTATGDKHYCYPHFTCAVDTENIRRVNDCRDIIQRMHLKQYELL 458
|||||
Db 301 LFURISTATGDKHYCYPHFTCAVDTENIRRVNDCRDIIQRMHLKQYELL 351
|||||
RESULT 3
US-09-952-680A-18
; Sequence 18, Application US/09952680A
; Publication No. US20030087239A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Marty
; APPLICANT: Epstein, David
; TITLE OF INVENTION: Target Activated Biosensor and Methods of Using Same
; FILE REFERENCE: 23239-501
; CURRENT APPLICATION NUMBER: US/09/952,680A
; CURRENT FILING DATE: 2001-09-13
; PRIOR FILING DATE: 2001-09-13
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-952-680A-18
Query Match 75.8%; Score 1819; DB 3; Length 381;
Best Local Similarity 95.0%; Pred. No. 2.2e-130; Indels 0; Gaps 0;
Matches 345; Conservative 9; Mismatches 9;
QY 96 KEARKVSRGIDRMRLDKRDLOQTHRLLLGAGESGKSTIVQMRLHVNFGNPNPEKKQK 155
|||||
Db 19 KERREANKKIEKQLQKERLAYKATHRLLLGAGESGKSTIVQMRLHVNFGNPNPEKKQK 78
|||||
QY 156 ILDIRKNVKDAIVTIIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSQEFDFHVK 215
|||||
Db 79 ILDIRKNVKDAIVTIIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSQEFDFHVK 138
|||||
QY 216 KLWDDGKACFERSNEYQLIDCAQYFLERIDSVLVDTPTDQDLRLCRVLTSGIFETR 275
|||||
Db 139 KLWDDGKACFERSNEYQLIDCAQYFLERIDSVLVDTPTDQDLRLCRVLTSGIFETR 198
|||||
QY 276 FOVDKVNHEMDVGGORDERRKWIQCFNDVTAIIVAACSSYNMVIREDNNTNRLRESLD 335
|||||
Db 199 FOVDKVNHEMDVGGORDERRKWIQCFNDVTAIIVAACSSYNMVIREDNNTNRLRESLD 258
|||||
QY 336 LFESIWNNRWLRITISILFLNKQDMLAEKVLGAGSKIEDYFPEYANYTVPEDATPDAGED 395
|||||
Db 259 LFESIWNNRWLRITISILFLNKQDMLAEKVLGAGSKIEDYFPEYANYTVPEDATPDAGED 318
|||||
QY 396 PKVTRAKFFIRDLFURISTATGDKHYCYPHFTCAVDTENIRRVNDCRDIIQRMHLKQY 455
|||||
Db 319 PKVTRAKFFIRDLFURISTATGDKHYCYPHFTCAVDTENIRRVNDCRDIIQRMHLKQY 378
|||||
QY 456 ELL 458
|||||
Db 379 ELL 381
|||||
RESULT 4
US-10-215-982-18
; Sequence 18, Application US/10215982
; Publication No. US20040219523A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Martin
; APPLICANT: Epstein, David
; APPLICANT: Hamaguchi, Nobuko
; APPLICANT: Kurz, Markus
; APPLICANT: Keefe, Tony
; APPLICANT: Wilson, Charles
; APPLICANT: Grate, Dilara
; APPLICANT: Marshall, Kristin
; APPLICANT: McCauley, Thomas
; APPLICANT: Kurz, Jeffrey
; TITLE OF INVENTION: NUCLEIC ACID SENSOR MOLECULES AND METHODS OF USING SAME
; FILE REFERENCE: 23239-501 CIP
; CURRENT APPLICATION NUMBER: US/10/215,982
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/232,454
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 09/952,680
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: 60/311,378
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/313,932
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/338,186
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/349,959
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 60/364,486
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/376,744
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/367,991
; PRIOR FILING DATE: 2002-03-25
```

```
; PRIOR APPLICATION NUMBER: 60/369,887
; PRIOR FILING DATE: 2002-04-04
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 372
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 18
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-215-982-18

Query Match      75.8%; Score 1819; DB 5; Length 381;
Best Local Similarity 95.0%; Pred. No. 2.2e-130;
Matches 345; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

Qy 96 KEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGESGKSTIVKQMRILHVGNFPEEKQK 155
Db 19 KERREANKKIEKQLKERLAYKATHRLLLGAGESGKSTIVKQMRILHVGNFPEEKQK 78
Qy 156 ILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFFYSQBFDDHVK 215
Db 79 ILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFFYSQBFDDHVK 138
Qy 216 KLWDEGVKACFERSNEYQLIDCAQYFLERIDSLSVDYPTDQDLRCRVLTSIGIFETR 275
Db 139 KLWDEGVKACFERSNEYQLIDCAQYFLERIDSLSVDYPTDQDLRCRVLTSIGIFETR 198
Qy 276 FOVDKVNFMFDVGGQDERRKWIQCNDVTAIYVAACSSYNMVIREDNNTNRLRESLD 335
Db 199 FOVDKVNFMFDVGGQDERRKWIQCNDVTAIYVAACSSYNMVIREDNNTNRLRESLD 258
Qy 336 LFESIWNRRWLRTISILFLNKQDMLAEKVLGSKIEDYFPEYANTVTPEDATPDAGED 395
Db 259 LFESIWNRRWLRTISILFLNKQDMLAEKVLGSKIEDYFPEYANTVTPEDATPDAGED 318
Qy 396 PKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDRCRDI IQRMHLKQY 455
Db 319 PKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDRCRDI IQRMHLKQY 378
Qy 456 ELL 458
Db 379 ELL 381

RESULT 6
US-10-732-923-8012
; Sequence 8012, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8012
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-732-923-8012

Query Match      75.8%; Score 1819; DB 5; Length 381;
Best Local Similarity 95.0%; Pred. No. 2.2e-130;
Matches 345; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

Qy 96 KEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGESGKSTIVKQMRILHVGNFPEEKQK 155
Db 19 KERREANKKIEKQLKERLAYKATHRLLLGAGESGKSTIVKQMRILHVGNFPEEKQK 78
Qy 156 ILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFFYSQBFDDHVK 215
Db 79 ILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFFYSQBFDDHVK 138
Qy 216 KLWDEGVKACFERSNEYQLIDCAQYFLERIDSLSVDYPTDQDLRCRVLTSIGIFETR 275
Db 139 KLWDEGVKACFERSNEYQLIDCAQYFLERIDSLSVDYPTDQDLRCRVLTSIGIFETR 198
Qy 276 FOVDKVNFMFDVGGQDERRKWIQCNDVTAIYVAACSSYNMVIREDNNTNRLRESLD 335
Db 199 FOVDKVNFMFDVGGQDERRKWIQCNDVTAIYVAACSSYNMVIREDNNTNRLRESLD 258
Qy 336 LFESIWNRRWLRTISILFLNKQDMLAEKVLGSKIEDYFPEYANTVTPEDATPDAGED 395
Db 259 LFESIWNRRWLRTISILFLNKQDMLAEKVLGSKIEDYFPEYANTVTPEDATPDAGED 318
Qy 396 PKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDRCRDI IQRMHLKQY 455
Db 319 PKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDRCRDI IQRMHLKQY 378
Qy 456 ELL 458
Db 379 ELL 381

RESULT 5
US-10-732-923-7615
; Sequence 7615, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 7615
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-732-923-7615

Query Match      75.8%; Score 1819; DB 5; Length 381;
Best Local Similarity 95.0%; Pred. No. 2.2e-130;
Matches 345; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

Qy 96 KEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGESGKSTIVKQMRILHVGNFPEEKQK 155
Db 19 KERREANKKIEKQLKERLAYKATHRLLLGAGESGKSTIVKQMRILHVGNFPEEKQK 78
Qy 156 ILDIRKNVKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFFYSQBFDDHVK 215
```

## RESULT 7

US-10-756-149-5241  
; Sequence 5241, Application US/10756149  
; Publication No. US20050181375A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND  
; FILE REFERENCE: file  
; CURRENT FILING DATE: 2004-01-12  
; NUMBER OF SEQ ID NOS: 5818  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5241  
; LENGTH: 381  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-756-149-5241

Query Match 75.8%; Score 1819; DB 5; Length 381;  
Best Local Similarity 95.0%; Pred. No. 2.2e-130;  
Matches 345; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY	96	KEARKVSRGIDRMRLDQKRDLOQTHRLLLLGAGESGKSTIVKQMRILHVNFGNPEKKQK	155
DB	19	KERREANKKIEKQKERLAYKATHRLLLLGAGESGKSTIVKQMRILHVNFGNPEKKQK	78
QY	156	ILDIRKNVKDAIVTVISAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSOEFFDHVK	215
DB	79	ILDIRKNVKDAIVTVISAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSOEFFDHVK	138
QY	216	KLWDDGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYPTDQDLLRCRVLTSIGIFETR	275
DB	139	KLWDDGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYPTDQDLLRCRVLTSIGIFETR	198
QY	276	FQVDKYNFHMFDVGGQDERRKWIQCFNDVTAIYVAACSSYNMVIREDNNTNRLRESLD	335
DB	199	FQVDKYNFHMFDVGGQDERRKWIQCFNDVTAIYVAACSSYNMVIREDNNTNRLRESLD	258
QY	336	LFESIWNRLRTISILFLNKQDMLAEKVLAKGSKIETYFPEYANTVPEDATPDAGED	395
DB	259	LFESIWNRLRTISILFLNKQDMLAEKVLAKGSKIETYFPEYANTVPEDATPDAGED	318
QY	396	PKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQY	455
DB	319	PKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQY	378
QY	456	ELL 458	
DB	379	ELL 381	

## RESULT 8

US-10-732-923-7651  
; Sequence 7651, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 7651  
; LENGTH: 381  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-732-923-7651

Query Match 75.5%; Score 1811; DB 5; Length 381;  
Best Local Similarity 94.8%; Pred. No. 8.9e-130;  
Matches 344; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY	96	KEARKVSRGIDRMRLDQKRDLOQTHRLLLLGAGESGKSTIVKQMRILHVNFGNPEKKQK	155
DB	19	KERREANKKIEKQKERLAYKATHRLLLLGAGESGKSTIVKQMRILHVNFGNPEKKQK	78
QY	156	ILDIRKNVKDAIVTVISAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSOEFFDHVK	215
DB	79	ILDIRKNVKDAIVTVISAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSOEFFDHVK	138
QY	216	KLWDDGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYPTDQDLLRCRVLTSIGIFETR	275
DB	139	KLWDDGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYPTDQDLLRCRVLTSIGIFETR	198
QY	276	FQVDKYNFHMFDVGGQDERRKWIQCFNDVTAIYVAACSSYNMVIREDNNTNRLRESLD	335
DB	199	FQVDKYNFHMFDVGGQDERRKWIQCFNDVTAIYVAACSSYNMVIREDNNTNRLRESLD	258
QY	336	LFESIWNRLRTISILFLNKQDMLAEKVLAKGSKIETYFPEYANTVPEDATPDAGED	395
DB	259	LFESIWNRLRTISILFLNKQDMLAEKVLAKGSKIETYFPEYANTVPEDATPDAGED	318
QY	396	PKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQY	455
DB	319	PKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDCRDIIQRMHLKQY	378
QY	456	ELL 458	
DB	379	ELL 381	

## RESULT 9

US-10-732-923-7963  
; Sequence 7963, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 7963  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Xenopus laevis  
US-10-732-923-7963

Query Match 69.9%; Score 1677; DB 5; Length 379;  
Best Local Similarity 86.2%; Pred. No. 1.5e-119;  
Matches 313; Conservative 25; Mismatches 25; Indels 0; Gaps 0;

QY	96	KEARKVSRGIDRMRLDQKRDLOQTHRLLLLGAGESGKSTIVKQMRILHVNFGNPEKKQK	155
DB	17	KAQREANKKIEKQKERLAYKATHRLLLLGAGESGKSTIVKQMRILHVNFGNPEKKQK	76
QY	156	ILDIRKNVKDAIVTVISAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSOEFFDHVK	215
DB	77	SQDIRKNVKDAIVTVISAMSTIIPVPLANPENQFRSDYIKSIAPITDPEYSOEFFDHVK	136
QY	216	KLWDDGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYPTDQDLLRCRVLTSIGIFETR	275
DB	137	KLWDDGVKACFERSNEYQLIDCAQYFLERIDSVSLVDYPTDQDLLRCRVLTSIGIFETR	196
QY	276	FQVDKYNFHMFDVGGQDERRKWIQCFNDVTAIYVAACSSYNMVIREDNNTNRLRESLD	335
DB	197	FQVDKYNFHMFDVGGQDERRKWIQCFNDVTAIYVAACSSYNMVIREDNNTNRLRESLD	256

QY 336 LFSIMNNRMLRTISIIILFNKQDMLAEKVLGKSKIEDYFPEYANTVPEDATPDAGED 395  
 DB 257 LFSIMNNRMLRTISIIILFNKQDMLAEKVLGKSKIEDYFPEYANTVPEDATPDAGED 316  
 QY 396 PKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNCDDIIQRMHLQY 455  
 DB 317 PKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNCDDIIQRMHLQY 376  
 QY 456 ELL 458  
 DB 377 ELL 379  
 RESULT 10  
 US-09-963-131-184  
 ; Sequence 184, Application US/09963131  
 ; Publication No. US20030224460A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pedersen, Finn Skou  
 ; APPLICANT: Sorensen, Annette Balle  
 ; APPLICANT: Hernandez, Javier Martin  
 ; APPLICANT: Nielsen, Anne Ahlmann  
 ; APPLICANT: Moving, Helle  
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR LYMPHOMA AND LEUKEMIA  
 ; FILE REFERENCE: 529452000323  
 ; CURRENT APPLICATION NUMBER: US/09/963,131  
 ; CURRENT FILING DATE: 2001-09-24  
 ; PRIOR APPLICATION NUMBER: US 09/668,644  
 ; PRIOR FILING DATE: 2000-09-22  
 ; PRIOR APPLICATION NUMBER: US 09/905,390  
 ; PRIOR FILING DATE: 2001-07-13  
 ; PRIOR APPLICATION NUMBER: US 09/905,491  
 ; PRIOR FILING DATE: 2001-07-13  
 ; NUMBER OF SEQ ID NOS: 215  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 184  
 ; LENGTH: 756  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-963-131-184  
 Query Match 65.0%; Score 1559; DB 3; Length 756;  
 Best Local Similarity 72.6%; Pred. No. 4e-110;  
 Matches 292; Conservative 48; Mismatches 60; Indels 2; Gaps 2;  
 QY 59 GSPACARPADKP-KEKQRTTEOLSAEEREAEREAVKVRSGIDRMLDQKRDQLQ 117  
 DB 355 GLSECTRSRLSPGAKDPMEERRKQMKAEIEMREKQKADKKRSLIDKQLEEEKMDYM 414  
 QY 118 QTHRLLLGAGESGKSTIVKQMRILHVGNGFNP-EKKQKILDIRKNVKDAIVTVSAMST 176  
 DB 415 CTHRLLLGAGESGKSTIVKQMRILHVGNGFNGDSEKATKQDINKNLKEAIEITIVAAMSN 474  
 QY 177 IIPPVPLANPENQFRSDYIKSIAPITDFEYSQBFHVKKLWDEGVKACFERSNEYQLI 236  
 DB 475 LVPPVELANPENQFRVDYILSVNVPNFPDFPFYEHAHALWEDEGVACRYERSNEYQLI 534  
 QY 237 DCAQYFLERIDSVLVDYTPDODLLRCRLVLTSGIFETRFQVDKVNPFHMFVGGQDERR 296  
 DB 535 DCAQYFLDKIDVIKQADYVPSDQDLRLCRVLTSGIFETRFQVDKVNPFHMFVGGQDERR 594  
 QY 297 KWIQCNDVTAIIVVAAACSSNNVIREDDNNTNRLRESLDLFSIWNRRMLRTISIIILFN 356  
 DB 595 KWIQCNDVTAIIVVAAACSSNNVIREDDNNTNRLQELALNLFKSIWNRRMLRTISIIILFN 654  
 QY 357 KQDLAEKVLGKSKIEDYFPEYANTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTAT 416  
 DB 655 KQDLAEKVLGKSKIEDYFPEYANTVPEDATPDAGEDPRVTRAKYFIRDEFIRISTAS 714  
 QY 417 GDGKHVCYPHFTCAVDTENIRRVFNCDDIIQRMHLQYELL 458  
 DB 715 GDGRHYCYPHFTCAVDTENIRRVFNCDDIIQRMHLQYELL 756  
 RESULT 11  
 US-10-732-923-7617  
 ; Sequence 7617, Application US/10732923  
 ; Publication No. US20050108791A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 7617  
 ; LENGTH: 756  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-10-732-923-7617  
 Query Match 65.0%; Score 1559; DB 5; Length 756;  
 Best Local Similarity 72.6%; Pred. No. 4e-110;  
 Matches 292; Conservative 48; Mismatches 60; Indels 2; Gaps 2;  
 QY 59 GSPACARPADKP-KEKQRTTEOLSAEEREAEREAVKVRSGIDRMLDQKRDQLQ 117  
 DB 355 GLSECTRSRLSPGAKDPMEERRKQMKAEIEMREKQKADKKRSLIDKQLEEEKMDYM 414  
 QY 118 QTHRLLLGAGESGKSTIVKQMRILHVGNGFNP-EKKQKILDIRKNVKDAIVTVSAMST 176  
 DB 415 CTHRLLLGAGESGKSTIVKQMRILHVGNGFNGDSEKATKQDINKNLKEAIEITIVAAMSN 474  
 QY 177 IIPPVPLANPENQFRSDYIKSIAPITDFEYSQBFHVKKLWDEGVKACFERSNEYQLI 236  
 DB 475 LVPPVELANPENQFRVDYILSVNVPNFPDFPFYEHAHALWEDEGVACRYERSNEYQLI 534  
 QY 237 DCAQYFLERIDSVLVDYTPDODLLRCRLVLTSGIFETRFQVDKVNPFHMFVGGQDERR 296  
 DB 535 DCAQYFLDKIDVIKQADYVPSDQDLRLCRVLTSGIFETRFQVDKVNPFHMFVGGQDERR 594  
 QY 297 KWIQCNDVTAIIVVAAACSSNNVIREDDNNTNRLRESLDLFSIWNRRMLRTISIIILFN 356  
 DB 595 KWIQCNDVTAIIVVAAACSSNNVIREDDNNTNRLQELALNLFKSIWNRRMLRTISIIILFN 654  
 QY 357 KQDLAEKVLGKSKIEDYFPEYANTVPEDATPDAGEDPKVTRAKFFIRDLFLRISTAT 416  
 DB 655 KQDLAEKVLGKSKIEDYFPEYANTVPEDATPDAGEDPRVTRAKYFIRDEFIRISTAS 714  
 QY 417 GDGKHVCYPHFTCAVDTENIRRVFNCDDIIQRMHLQYELL 458  
 DB 715 GDGRHYCYPHFTCAVDTENIRRVFNCDDIIQRMHLQYELL 756  
 RESULT 12  
 US-10-732-923-8020  
 ; Sequence 8020, Application US/10732923  
 ; Publication No. US20050108791A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 8020  
 ; LENGTH: 909  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-732-923-8020  
 Query Match 65.0%; Score 1559; DB 5; Length 909;

Best Local Similarity 68.6%; Pred. No. 5.1e-110;  
Matches 300; Conservative 49; Mismatches 68; Indels 20; Gaps 4;

QY 42 RAAARDARTLLPR---GGGSPACARPADKP-KEKQRTQTEQLSAREEAAKEREAVKE 97  
DB 473 RKPQRLNRLVQAEGCGFGRSESPQPKASRLKVKVPLAEKRRQMRKEALEKRAQKR 532  
QY 98 A-BKVSRGIDRLMDQKRDILQOQTHRLLLGAGSGSGKSTIVKQMRILHVGNGFNP-----150  
DB 533 AEKRSKLIDQJQDEKQKGYMCTHRLLLGAGSGSGKSTIVKQMRILHVGNGFNGEGEEDP 592  
QY 151 -----EKKQKILDIRKNVDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAP 201  
DB 593 QAARSNSDGEKATKVQDQIKNNLKEAETIIVAASNLVPPVELANPENQFRVDYILSVNV 652  
QY 202 TDFEYSGEFPDVKVKKWDDGKACAFERSNEYQLIDCAQYFLERIDSLSVLTPTDQDL 261  
DB 653 PDFDFPEFEYEHAKALWEDEGRVACYSNEYQLIDCAQYFLDKIDVIKQADYVPSDQDL 712  
QY 262 LRCRVLTSGIFETRFOVDKYNFHMEDVGGORDERRKWIQCFNDVTALIIYVAACSSYNMVI 321  
DB 713 LRCRVLTSGIFETRFOVDKYNFHMEDVGGORDERRKWIQCFNDVTALIIYVAACSSYNMVI 772  
QY 322 REDNTNRLRESLDFESIWNRLRTISILFLNKQDMLAEKVLGAGSKIEDYFFPEYAN 381  
DB 773 REDNTNRLQALNLFKSIWNNRLRTISILFLNKQDMLAEKVLGAGSKIEDYFFPEFAR 832  
QY 382 YTPEDATPDAGDPKVTAKFIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFN 441  
DB 833 YTPEDATPDGDPKVTAKFIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFN 892  
QY 442 DCRDIIQRMHLQYELL 458  
DB 893 DCRDIIQRMHLQYELL 909

RESULT 13  
US-10-732-923-7654  
; Sequence 7654, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 7654  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-732-923-7654

Query Match 64.5%; Score 1548; DB 5; Length 715;  
Best Local Similarity 70.2%; Pred. No. 2.5e-109;  
Matches 292; Conservative 47; Mismatches 61; Indels 16; Gaps 2;

QY 59 GSPACARPADKP-KEKQRTQTEQLSAREEAAKEREAVKEARVSRGIDRLMDQKRDQ 117  
DB 300 GLSECTRSLSFGKADPMEERRKQMRKEAMREKQKADKRSKLIDKQLEEKWDYM 359  
QY 118 QTHRLLLGAGSGSGKSTIVKQMRILHVGNGFNP-----EKKQKILDIRKN 162  
DB 360 CTHRLLLGAGSGSGKSTIVKQMRILHVGNGFNGEGEEDPQAARSNSDGEKATKVQDIKNN 419  
QY 163 VKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQEFFDHVKKLWDEG 222  
DB 420 LKEAETIVAASNLVPPVELANPENQFRVDYILSVNVNPNFPPPEFEYEHAKALWEDEG 479  
QY 223 VKACFERSNEYQLIDCAQYFLERIDSLSVLTPTDQDLRRCRVLTSGIFETRFOVDKVN 282

Best Local Similarity 68.6%; Pred. No. 5.1e-110;  
Matches 300; Conservative 49; Mismatches 68; Indels 20; Gaps 4;

QY 42 RAAARDARTLLPR---GGGSPACARPADKP-KEKQRTQTEQLSAREEAAKEREAVKE 97  
DB 473 RKPQRLNRLVQAEGCGFGRSESPQPKASRLKVKVPLAEKRRQMRKEALEKRAQKR 532  
QY 98 A-BKVSRGIDRLMDQKRDILQOQTHRLLLGAGSGSGKSTIVKQMRILHVGNGFNP-----150  
DB 533 AEKRSKLIDQJQDEKQKGYMCTHRLLLGAGSGSGKSTIVKQMRILHVGNGFNGEGEEDP 592  
QY 151 -----EKKQKILDIRKNVDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAP 201  
DB 593 QAARSNSDGEKATKVQDQIKNNLKEAETIIVAASNLVPPVELANPENQFRVDYILSVNV 652  
QY 202 TDFEYSGEFPDVKVKKWDDGKACAFERSNEYQLIDCAQYFLERIDSLSVLTPTDQDL 261  
DB 653 PDFDFPEFEYEHAKALWEDEGRVACYSNEYQLIDCAQYFLDKIDVIKQADYVPSDQDL 712  
QY 262 LRCRVLTSGIFETRFOVDKYNFHMEDVGGORDERRKWIQCFNDVTALIIYVAACSSYNMVI 321  
DB 713 LRCRVLTSGIFETRFOVDKYNFHMEDVGGORDERRKWIQCFNDVTALIIYVAACSSYNMVI 772  
QY 322 REDNTNRLRESLDFESIWNRLRTISILFLNKQDMLAEKVLGAGSKIEDYFFPEYAN 381  
DB 773 REDNTNRLQALNLFKSIWNNRLRTISILFLNKQDMLAEKVLGAGSKIEDYFFPEFAR 832  
QY 382 YTPEDATPDAGDPKVTAKFIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFN 441  
DB 833 YTPEDATPDGDPKVTAKFIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFN 892  
QY 442 DCRDIIQRMHLQYELL 458  
DB 893 DCRDIIQRMHLQYELL 909

RESULT 13  
US-10-732-923-7654  
; Sequence 7654, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 7654  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-732-923-7654

Query Match 64.5%; Score 1548; DB 5; Length 715;  
Best Local Similarity 70.2%; Pred. No. 2.5e-109;  
Matches 292; Conservative 47; Mismatches 61; Indels 16; Gaps 2;

QY 59 GSPACARPADKP-KEKQRTQTEQLSAREEAAKEREAVKEARVSRGIDRLMDQKRDQ 117  
DB 300 GLSECTRSLSFGKADPMEERRKQMRKEAMREKQKADKRSKLIDKQLEEKWDYM 359  
QY 118 QTHRLLLGAGSGSGKSTIVKQMRILHVGNGFNP-----EKKQKILDIRKN 162  
DB 360 CTHRLLLGAGSGSGKSTIVKQMRILHVGNGFNGEGEEDPQAARSNSDGEKATKVQDIKNN 419  
QY 163 VKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQEFFDHVKKLWDEG 222  
DB 420 LKEAETIVAASNLVPPVELANPENQFRVDYILSVNVNPNFPPPEFEYEHAKALWEDEG 479  
QY 223 VKACFERSNEYQLIDCAQYFLERIDSLSVLTPTDQDLRRCRVLTSGIFETRFOVDKVN 282

DB 480 VRACYSRNEYQLIDCAQYFLDKIDVIKQADYVPSDQDLPRCRVLTSGIFETRFOVDKVN 539  
QY 283 FHMFDVGGORDERRKWIQCFNDVTALIIYVAACSSYNMVIREDNTNRLRESLDFESIWN 342  
DB 540 FHMFDVGGORDERRKWIQCFNDVTALIIYVAACSSYNMVIREDNTNRLRESLDFESIWN 599  
QY 343 NWLRTISILFLNKQDMLAEKVLGAGSKIEDYFFPEYANYPEDATPDAGDPKVTAK 402  
DB 600 NWLRTISILFLNKQDMLAEKVLGAGSKIEDYFFPEYANYPEDATPDAGDPKVTAK 659  
QY 403 FTIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFNDCRDIIOQMHLQYELL 458  
DB 791 YFIRDFLRLSTASGDRHYCYPHFTCAVDTENIRRVFNDCRDIIOQMHLQYELL 715

RESULT 14  
US-10-732-923-7655  
; Sequence 7655, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 7655  
; LENGTH: 846  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-732-923-7655

Query Match 64.5%; Score 1548; DB 5; Length 846;  
Best Local Similarity 70.2%; Pred. No. 3.2e-109;  
Matches 292; Conservative 47; Mismatches 61; Indels 16; Gaps 2;

QY 59 GSPACARPADKP-KEKQRTQTEQLSAREEAAKEREAVKEARVSRGIDRLMDQKRDQ 117  
DB 431 GLSECTRSLSFGKADPMEERRKQMRKEAMREKQKADKRSKLIDKQLEEKWDYM 490  
QY 118 QTHRLLLGAGSGSGKSTIVKQMRILHVGNGFNP-----EKKQKILDIRKN 162  
DB 491 CTHRLLLGAGSGSGKSTIVKQMRILHVGNGFNGEGEEDPQAARSNSDGEKATKVQDIKNN 550  
QY 163 VKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQEFFDHVKKLWDEG 222  
DB 551 LKEAETIVAASNLVPPVELANPENQFRVDYILSVNVNPNFPPPEFEYEHAKALWEDEG 610  
QY 223 VKACFERSNEYQLIDCAQYFLERIDSLSVLTPTDQDLRRCRVLTSGIFETRFOVDKVN 282  
DB 611 VRACYSRNEYQLIDCAQYFLDKIDVIKQADYVPSDQDLPRCRVLTSGIFETRFOVDKVN 670  
QY 283 FHMFDVGGORDERRKWIQCFNDVTALIIYVAACSSYNMVIREDNTNRLRESLDFESIWN 342  
DB 671 FHMFDVGGORDERRKWIQCFNDVTALIIYVAACSSYNMVIREDNTNRLRESLDFESIWN 730  
QY 343 NWLRTISILFLNKQDMLAEKVLGAGSKIEDYFFPEYANYPEDATPDAGDPKVTAK 402  
DB 731 NWLRTISILFLNKQDMLAEKVLGAGSKIEDYFFPEYANYPEDATPDAGDPKVTAK 790  
QY 403 FTIRDLFLRISTATGDKHYCYPHFTCAVDTENIRRVFNDCRDIIOQMHLQYELL 458  
DB 791 YFIRDFLRLSTASGDRHYCYPHFTCAVDTENIRRVFNDCRDIIOQMHLQYELL 846

RESULT 15  
US-10-732-923-7656  
; Sequence 7656, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D



; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

; FILE REFERENCE: 38-15(52796)C

; CURRENT APPLICATION NUMBER: US/10/732,923

; PRIOR FILING DATE: 2003-12-10

; PRIOR APPLICATION NUMBER: 10/310,154

; PRIOR FILING DATE: 2002-12-04

; NUMBER OF SEQ ID NOS: 24149

; SEQ ID NO 7656

; LENGTH: 846

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-10-732-923-7656

Query Match 64.2%; Score 1541; DB 5; Length 846;

Best Local Similarity 70.0%; Pred. No. 1.1e-108;

Matches 291; Conservative 47; Mismatches 62; Indels 16; Gaps 2;

```
Qy 59 GSPACARPKADKP-KEKQRTEQLSAEREAREAKVSRGIDRMRLRQKQDLQ 117
Db 431 GLSECTSRSLSPGAKDPMERKKQMRKEAWEMREKQKADKKRKLIDKQLEEEKMDYM 490
Qy 118 QTHRLLLIGAGESKSTIVKQMRILHVNGFNPE-----EKKQKILDIRKN 162
Db 491 CTHRLLLIGAGESKSTIVKQMRILHVNGFNGEDEEDPQAARSNSDGEKATKVQDIKN 550
Qy 163 VKDAIVTIVSAMSTIIPVPLANPENQFRSDYIKSIAPITDFEYSQBFDDHVKKLWDEG 222
Db 551 LKEAETIVAMSNLPPVELANPENQFRVDYILSVNVPNFPDFPEFYEHAHALWEDEG 610
Qy 223 VKACFERSNEYQLIDCAQYFLERIDSVSLVDYTTDODLLRCRVLTSGIPETRFOVDKN 282
Db 611 VRACYERSNEYQLIDCAWFUDKIDVIKQADYVPSDQDLPCRVLTSIGIPETKFOVDKN 670
Qy 283 FHMEDVGQDERRKWKIQCNDVTAIYVAACSSYNMVIREDNNTNRLRESLDLFSIWN 342
Db 671 FHMEDVGQDERRKWKIQCNDVTAIIFVVAASSYNMVIREDNQTNRLOEALNLFKSIWN 730
Qy 343 NRWLRTISILFLNKKQDLAEKVLGSKIEDYFPEYANYTVPEDATPDAGEDPKVTRAK 402
Db 731 NRWLRTISVILFNKQDLAEKVLGSKIEDYFPEFARYTTPEDATPDAGEDPRVTRAK 790
Qy 403 FFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDCRDIIOHMLKQYELL 458
Db 791 YFIRDFLRISTASGDGRHYCYPHFTCAVDTENIRRVFNDCRDIIOHMLRQYELL 846
```

Search completed: March 2, 2006, 19:41:16

Job time : 210.863 secs

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein --protein search, using sw model

Run on: March 2, 2006, 19:37:27 ; Search time 15.9223 Seconds  
(without alignments)  
575.298 Million cell updates/sec

Title: US-10-618-320A-1

Perfect score: 2400

Sequence: 1 MGLCYSLRLLFGPGDDPC.....VFNDCRDIQRMHMKOYELL 458

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 135339 seqs, 20000136 residues

Total number of hits satisfying chosen parameters: 135339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

D. Database : Published Applications AA\_New.\*

- 1: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pap.\*
- 2: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pap.\*
- 3: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pap.\*
- 4: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pap.\*
- 5: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pap.\*
- 6: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pap.\*
- 7: /cgn2\_6/ptodata/2/pubppaa/US11\_NEW\_PUB.pap.\*
- 8: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2400	100.0	458	6	US-10-618-320A-1
2	2124	88.5	448	6	US-10-618-320A-25
3	2113	88.0	450	6	US-10-618-320A-26
4	1538	64.1	864	7	US-11-053-100-58
5	1525.5	63.6	394	6	US-10-821-234-1626
6	728.5	30.4	353	7	US-11-060-023-4
7	728.5	30.4	353	7	US-11-060-023-8
8	723.5	30.1	859	7	US-11-053-100-52
9	723.5	30.1	1309	7	US-11-053-100-53
10	720.5	30.0	353	7	US-11-060-023-15
11	719.5	30.0	359	7	US-11-060-023-13
12	717.5	29.9	353	7	US-11-060-023-12
13	717.5	29.9	353	7	US-11-060-023-14
14	717.5	29.9	359	7	US-11-060-023-11
15	717.5	29.9	360	7	US-11-060-023-17
16	716.5	29.9	353	7	US-11-060-023-16
17	705.5	29.4	353	7	US-11-060-023-2
18	705.5	29.4	353	7	US-11-060-023-6
19	691.5	28.8	408	6	US-10-821-234-1100
20	672.5	28.0	374	7	US-11-060-023-10
21	672.5	28.0	374	7	US-11-228-364-2
22	672.5	28.0	374	7	US-11-159-041-188
23	647.5	27.0	374	7	US-11-228-364-4
24	125.5	5.2	627	7	US-11-150-845-47
25	122	5.1	179	7	US-11-093-746A-25

#### ALIGNMENTS

##### RESULT 1

US-10-618-320A-1

; Sequence 1, Application US/10618320A

; Publication No. US20050260595A1

; GENERAL INFORMATION: Sumitomo Chemical Company Limited

; TITLE OF INVENTION: NOVEL G PROTEINS, POLYNUCLEOTIDE ENCODING THE SAME AND UTILIZATION

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/618,320A

; CURRENT FILING DATE: 2003-07-11

; PRIOR APPLICATION NUMBER: JP 2002/206841

; PRIOR FILING DATE: 2002-07-16

; PRIOR APPLICATION NUMBER: JP 2002/367778

; PRIOR FILING DATE: 2002-12-19

; PRIOR APPLICATION NUMBER: JP 2003/095955

; PRIOR FILING DATE: 2003-03-31

; NUMBER OF SEQ ID NOS: 34

; SEQ ID NO 1:

; LENGTH: 458

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-618-320A-1

Query Match 100.0%; Score 2400; DB 6; Length 458;

Best Local Similarity 100.0%; Pred. No. 1.4e-170;

Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLCYSLRLLFGPGDDPCAASEPPVEDAQPAPALAPVRAAARDTARTILPRGEGS 60

DB 1 MGLCYSLRLLFGPGDDPCAASEPPVEDAQPAPALAPVRAAARDTARTILPRGEGS 60

QY 61 PACARPKADKPKRKRTQTEQLSAEREAKEREAKVSGIDRMRLDQRDLQOOTH 120

DB 61 PACARPKADKPKRKRTQTEQLSAEREAKEREAKVSGIDRMRLDQRDLQOOTH 120

QY 121 RLLLAGSGSGSTIVKQMRILHVNGFNPEEKQKILDIRKNVKDAIVTIVSAMSTIPP 180

DB 121 RLLLAGSGSGSTIVKQMRILHVNGFNPEEKQKILDIRKNVKDAIVTIVSAMSTIPP 180

QY 181 VPLANPENFRSDYIISKIAPITDFFYSQBFDFHVKLWDEGVKACFERSNEVLIDCAQ 240

DB 181 VPLANPENFRSDYIISKIAPITDFFYSQBFDFHVKLWDEGVKACFERSNEVLIDCAQ 240

QY 241 YFLERIDSVSLVDYTTPTDQDLRCRVLTSIGIFRQVDKVNFMFVGVGQRRERKWIQ 300

DB 241 YFLERIDSVSLVDYTTPTDQDLRCRVLTSIGIFRQVDKVNFMFVGVGQRRERKWIQ 300

QY 301 CFNDVTATIIYVAACSSYNMVIREDNNTNRLRESLDLFEISWNNRNLRTISILFLNKQDM 360

DB 301 CFNDVTATIIYVAACSSYNMVIREDNNTNRLRESLDLFEISWNNRNLRTISILFLNKQDM 360

QY 360 CFNDVTATIIYVAACSSYNMVIREDNNTNRLRESLDLFEISWNNRNLRTISILFLNKQDM 360

DB 360 CFNDVTATIIYVAACSSYNMVIREDNNTNRLRESLDLFEISWNNRNLRTISILFLNKQDM 360

301	CFNDVTAIIVYAACSNNVIREDNNTNRLRESLDLFESIWNRRWLRTISIIILFLNQDM	360
361	LAEKVLAGSKIEDYFPPEVANYTPEDATPDAGBDPKVTRAKFFIRDLFLRISTATGDCK	420
361	LAEKVLAGSKIEDYFPPEVANYTPEDATPDAGBDPKVTRAKFFIRDLFLRISTATGDCK	420
421	HYCPHFTCAVDTENIRRVFNDCRDIIQRMHLKOYELL	458
421	HYCPHFTCAVDTENIRRVFNDCRDIIQRMHLKOYELL	458

## RESULT 2

```

US-10-618-320A-25
; Sequence 25, Application US/10618320A
; Publication No. US2005026095A1
; GENERAL INFORMATION:
; APPLICANT: Sumitomo Chemical Company Limited
; TITLE OF INVENTION: NOVEL G PROTEINS, POLYNUCLEOTIDE ENCODING THE SAME AND UTILIZATION THEREOF
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/618,320A
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: JP 2002/206841
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: JP 2002/367778
; PRIOR FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: JP 2003/095955
; PRIOR FILING DATE: 2003-03-31
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 25
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-618-320A-25

```

### RESULT 3

```

US-10-618-320A-26
; Sequence 26, Application US/10618320A
; Publication No. US20050260595A1
; GENERAL INFORMATION:
; APPLICANT: Sumitomo Chemical Company Limited
; TITLE OF INVENTION: NOVEL G PROTEINS, POLYNUCLEOTIDE ENCODING THE SAME AND
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/618,320A
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: JP 2002/206841
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: JP 2002/367778
; PRIOR FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: JP 2003/095955
; PRIOR FILING DATE: 2003-03-31
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 26
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-618-320A-26

```

## RESULT 4

```

US-11-053-100-58
; Sequence 58, Application US/11053100
; Publication No. US2005025554A1
; GENERAL INFORMATION:
; APPLICANT: CHILKOTI, Ashutosh
; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION
; FILE REFERENCE: 4176-101 CIP
; CURRENT APPLICATION NUMBER: US/11/053,100
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US 09/812,382
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,659
; PRIOR FILING DATE: 2000-03-20

```

```
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 58
; LENGTH: 864
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(864)
; OTHER INFORMATION: pET15b-SD3-ELP1-90-throm-G protein alpha S
US-11-053-100-58

Query Match          64.1%; Score 1538; DB 7; Length 864;
Best Local Similarity 70.8%; Pred. No. 2.4e-106;
Matches 289; Conservative 47; Mismatches 56; Indels 16; Gaps 2;

QY 52 LLPRGEGSPACAPKADPKPKRQRTQQLSAEREAREAVKEARKVSRGIDRMRLD 111
DB 472 LVPRGSHMPWALEMCGCLGNSKTDQNEE-----KAQREANKKIEKOLQK 516
QY 112 QKRDLOQTHRLLLGAGESGKSTIVKQMRILHVNPNP-EKKQKILDIRKNVKDAIVTI 170
DB 517 DKQVYRATHRLLLGAGESGKSTIVKQMRILHVNPNP-SEKATKQDIKNLKEAETII 576
QY 171 VSMSTIIPVPLANPENFRSYIKSIAPITDFYSYQSFDFHVKLWDDGKACFERS 230
DB 577 VAAMSNLVPVELANPENFRVDYILSVNVPDFPPEFYEHAKEALWEDGVRACYS 636
QY 231 NEVQLDCAQYFLERIDSVLDVYTTDQDLRCRVLTSGIFPSTRQVDKVNFMFDVGG 290
DB 637 NEVQLDCAQYFLDKIDVQADYVPSDQDLRCRVLTSGIFETKPFQVDKVNFMFDVGG 696
QY 291 QDERERKVIQCFNDVTAIVAAACSSVMVIREDDNNTNRLRESLDLFSIWNRLRTIS 350
DB 697 QDERERKVIQCFNDVTAIVFAVASSVMVIREDDNNTNRLRESLDLFSIWNRLRTIS 756
QY 351 IILFLNKQDLAEKVLGSKIEDYFPEYANYTVPEADPDAGEDPKVTRAKFFIRDLFL 410
DB 757 VILFLNKQDLAEKVLGSKIEDYFPEFARYTTPEDATPECGEDPRVTRAKYFIRDEFL 816
QY 411 RISTATGDKGHYCPHFTCAVDTENIRRVFNDCRDIIQRMHLKQYELL 458
DB 817 RISTASGDRGHYCPHFTCAVDTENIRRVFNDCRDIIQRMHLKQYELL 864

RESULT 5
US-10-821-234-1626
; Sequence 1626, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1626
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1626

Query Match          53.6%; Score 1525.5; DB 6; Length 394;
Best Local Similarity 73.8%; Pred. No. 7.4e-106;
Matches 282; Conservative 44; Mismatches 42; Indels 15; Gaps 1;

QY 91 EREAVKEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGESGKSTIVKQMRILHVNPNP 150
DB 12 QRENEKAQREANKKIEKOLQKQVYRATHRLLLGAGESGKSTIVKQMRILHVNPNP 71
QY 151 -----EKKQKILDIRKNVKDAIVTIIVSAMSTIIPVPLANPENQFRSDYI 195
DB 72 GSEEDPQAARSNSDGEKATKQDIKNLKEAETIVAAAMSNLVPVELANPENQFRVDYI 131
QY 196 KSIAPITDFYSYQSFDFHVKLWDDGKACFERSNEYQLIDCAQYFLERIDSVLDVYT 255
DB 132 LSVNVPDFPPEFYEHAKEALWEDGVRACYSNEYQLIDCAQYFLDKIDVQADYV 191
QY 256 PTDQDLRCRVLTSGIFPSTRQVDKVNFMFDVGGQDERRKKIQCNDVTAIVAAAC 315
DB 192 PSDQDLRCRVLTSGIFETKPFQVDKVNFMFDVGGQDERRKKIQCNDVTAIVFAV 251
QY 316 SYNVMVIREDDNNTNRLRESLDLFSIWNRLRTISILFLNKQDLAEKVLGSKIEDY 375
DB 252 SYNVMVIREDDNNTNRLRESLDLFSIWNRLRTISILFLNKQDLAEKVLGSKIEDY 311
QY 376 FPEYANYTVPEADPDAGEDPKVTRAKFFIRDLFLRISTATGDKGHYCPHFTCAVD 435
DB 312 FPEFARYTTPEDATPECGEDPRVTRAKYFIRDEFLRISTASGDRGHYCPHFTCAVD 371
QY 436 IRRVFNDCRDIQRMHLKQYELL 458
DB 372 IRRVFNDCRDIQRMHLKQYELL 394

RESULT 6
US-11-060-023-4
; Sequence 4, Application US/11060023
; Publication No. US2005025531A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pharma Deutschland GmbH
; TITLE OF INVENTION: Process for identifying modulators of G protein coupled
; TITLE OF INVENTION: receptors
; FILE REFERENCE: AVE D-2000/A033 english
; CURRENT APPLICATION NUMBER: US/11/060,023
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: US/09/899,295
; PRIOR FILING DATE: 2003-07-06
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-060-023-4

Query Match          30.4%; Score 728.5; DB 7; Length 353;
Best Local Similarity 41.3%; Pred. No. 7.5e-47;
Matches 151; Conservative 69; Mismatches 127; Indels 19; Gaps 4;

QY 93 EAVKEARKVSRGIDRMRLDQKRDLOQTHRLLLGAGESGKSTIVKQMRILHVNPNP 152
DB 7 EBAKEARRINDEIERHVRDKDARRELKLLLTGSGSKSTFIKQMRIHSGSYDEDK 66
QY 153 KOKILDIRKNVKDAIVTIIVSAMSTIIPVPLANPENQFRSDYIKSTAPITDFYSQ 212
DB 67 RGFTKLVIYQNIPTAMQAMIRAMDTL--KIPYKYEHNKAHAQLVREVDVEKVSAPEN 124
QY 213 HVKLLWDDGKACFERSNEYQLIDCAQYFLERIDSVLDVYTPTDQDLRCRVLTSGIF 272
DB 125 AIKSLWDDGKACFERSNEYQLIDCAQYFLERIDSVLDVYTPTDQDLRCRVLTSGIF 184
QY 273 ETRFQVDKVNFMFDVGGQDERRKKIQCNDVTAIVAAACSSVMVIREDDNNTNRL 332
DB 185 EYFPDLQSVIFRMVDVGGQDERRKKIQCNDVTAIVAAACSSVMVIREDDNNTNRL 244
QY 333 SLDLFSIWNRLRTISILFLNKQDLAEKVLGSKIEDYFPEYANYTVPEADTPDA 392
DB 185 EYFPDLQSVIFRMVDVGGQDERRKKIQCNDVTAIVAAACSSVMVIREDDNNTNRL 244
```

Db 245 SKALFRTIITYPWFQSSVILFLNKKDLLEEKIM--YSHLVDFPEY-----D 290

Qy 393 GEDPKVTRAKFFRDLFLAISTATGKHVCYPHFTCAVDTENIRRVNDCRDIORMHL 452

Db 291 GPORDAQAAREFLKMFVDLNP---DSDKIYSHFTCATDTENIRVFVAVKDTILQLNL 347

Qy 453 KQYELL 458

Db 348 KQYELL 353

RESULT 7

US-11-060-023-8

; Sequence 8, Application US/11060023

; Publication No. US20050255531A1

; GENERAL INFORMATION:

; APPLICANT: Aventis Pharma Deutschland GmbH

; TITLE OF INVENTION: Process for identifying modulators of G protein coupled

; FILE REFERENCE: AVE D-2000/A033 englisch

; CURRENT APPLICATION NUMBER: US/11/060,023

; CURRENT FILING DATE: 2005-02-17

; PRIOR FILING DATE: 2003-07-06

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 353

; TYPE: PRT

; ORGANISM: Mus musculus

US-11-060-023-8

Query Match 30.4%; Score 728.5; DB 7; Length 353;

Best Local Similarity 41.3%; Pred. No. 7.5e-47;

Matches 151; Conservative 69; Mismatches 127; Indels 19; Gaps 4;

Qy 93 EAVKEARKVSRGIDRMRLDQKRDLOQTHRLLLLGAGESGKSTIVKQMRILVHNGFNPEEK 152

Db 7 EAKKARRINDEIERHVRDKRDARRELKLLLTGSGSKSTFIKQMRIHSGYSDEK 66

Qy 153 KOKILDIRKNVDAIYTVISAMSTIIPPVLANPENQFRSDYKSIAPITDFEYSQEFPD 212

Db 67 RGFTKLAVYQNIIFTAMQAMIRAMDTL--KIPYKYEHNKAHAQLVREVDKVSFAFENPYVD 124

Qy 213 HVKLLWDDGKACFRSEYQLIDCAQYFLERIDSLSVLTPTDQDLLRCRVLTSGIF 272

Db 125 AIKSLWNDFGIQECYDRREYQLSDSTKYLNLDLRVADPAVLTQDVLVRVPTTGI 184

Qy 273 ETRFOVDKYNFMDFVGQDERRKWIQCFNDVTAIYVAAACSSYNNMIRENNTNRLRE 332

Db 185 EYFDFLQSVIFRMVDVGGQSRERKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 244

Qy 333 SLDLPESINNRWLRTISILFLNKQDMLAEKVLGAKSKIYDFPEYANYTVPEDATPDA 392

Db 245 SKALFRTIITYPWFQSSVILFLNKKDLLEEKIM--YSHLVDFPEY-----D 290

Qy 393 GEDPKVTRAKFFRDLFLAISTATGKHVCYPHFTCAVDTENIRRVNDCRDIORMHL 452

Db 291 GPORDAQAAREFLKMFVDLNP---DSDKIYSHFTCATDTENIRVFVAVKDTILQLNL 347

Qy 453 KQYELL 458

Db 348 KQYELL 353

RESULT 8

US-11-053-100-52

; Sequence 52, Application US/11053100

; Publication No. US20050255534A1

; GENERAL INFORMATION:

; APPLICANT: CHILKOTI, Ashutosh

; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION

; FILE REFERENCE: 4176-101 CIP

; CURRENT APPLICATION NUMBER: US/11/053,100

; CURRENT FILING DATE: 2005-02-08

; PRIOR APPLICATION NUMBER: US 09/812,382

; PRIOR FILING DATE: 2001-03-20

; PRIOR APPLICATION NUMBER: US 60/190,659

; PRIOR FILING DATE: 2000-03-20

; NUMBER OF SEQ ID NOS: 58

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 52

; LENGTH: 859

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: Synthetic Construct

; FEATURE: MISC\_FEATURE

; NAME/KEY: MISC\_FEATURE

; LOCATION: (1)..(859)

; OTHER INFORMATION: pET15b-SD1-ELP1-90-throm-G protein alpha Q

US-11-053-100-52

Query Match 30.1%; Score 723.5; DB 7; Length 859;

Best Local Similarity 38.6%; Pred. No. 5.5e-46;

Matches 158; Conservative 76; Mismatches 154; Indels 21; Gaps 5;

Qy 52 LLPRGEGSPACARPKADPK--EXRQRTQQLSASEEAAKEREAVKEARKVSRGIDRML 109

Db 470 LVPRGSHMGLNDIFEAKQIEWHHPMALENTLSINACCLSEAKEARRINDEIERQL 529

Qy 110 RDQKRDLOQTHRLLLLGAGESGKSTIVKQMRILVHNGFNPEEKOKILDIRKNVDAIYTV 169

Db 530 RDKERDARRELKLLLTGSGSKSTFIKQMRIHSGYSDEKRGFTKLAVYQNIIFTAMQA 589

Qy 170 IVSAMSTIIPPVLANPENQFRSDYKSIAPITDFEYSQEFPDHVKLLWDDGKACFR 229

Db 590 MIRAMDTL--KIPYKYEHNKAHAQLVREVDKVSFAFENPYVDATKSLWNDFGIQECYDR 647

Qy 230 SNEYQLIDCAQYFLERIDSLSVLTPTDQDLLRCRVLTSGIFETRFQVDKYNFMDFVG 289

Db 648 RREYQLSSTKYLNLDLRVADPAVLTQDVLVRVPTTGIIEYFPDLQSVIFRMVDVG 707

Qy 290 GQDERRKWIQCFNDVTAIYVAAACSSYNNMIRENNTNRLRESLDLPESINNRWLRTI 349

Db 708 QQRSEKRWIHCENVTSIMFLVALSEYDQVLVESDNENRMEESKALFRTIITYPWFQNS 767

Qy 350 SLDLPESINNRWLRTISILFLNKQDMLAEKVLGAKSKIYDFPEYANYTVPEDATPDA 409

Db 768 SVILFLNKKDLLEEKIM--YSHLVDFPEY-----DGPORDAQAAREFLKMF 813

Qy 410 LRISTATGKHVCYPHFTCAVDTENIRRVNDCRDIORMHLKQYELL 458

Db 814 VDLNP---DSDKIYSHFTCATDTENIRVFVAVKDTILQLNLKEYNLV 859

RESULT 9

US-11-053-100-53

; Sequence 53, Application US/11053100

; Publication No. US20050255534A1

; GENERAL INFORMATION:

; APPLICANT: CHILKOTI, Ashutosh

; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION

; FILE REFERENCE: 4176-101 CIP

; CURRENT APPLICATION NUMBER: US/11/053,100

; CURRENT FILING DATE: 2005-02-08

; PRIOR APPLICATION NUMBER: US 09/812,382

; PRIOR FILING DATE: 2001-03-20

; PRIOR APPLICATION NUMBER: US 60/190,659

; PRIOR FILING DATE: 2000-03-20

; NUMBER OF SEQ ID NOS: 58

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 53

; LENGTH: 1309

; TYPE: PRT

; ORGANISM: Artificial







Db 131 AIKSLWDPGIECVDRRREYQLSDSTKYKYLNDLDRVADPAYLPTQDVLVRVPTTGII 190  
Qy 273 ETRFQVDKYNHMFVDCGQDRERKWIQCFNDVTAIYVACSSYNMVIREDNNTNRLRE 332  
Db 191 EYPDLQSVIFRMVDVGQSRERKWIHCFENVTSIMFLVALSEYDQVLVESDNENRMBE 250  
Qy 333 SLDLFESIMNNRWLRTISIILFNKQDMLAEKVLGKAGKIEDYPEYANYTVPEDATPDA 392  
Db 251 SKALFRITIIYFPWFQNSSVILFNKKDLLEKIM--YSHLVDYFPEY-----D 296  
Qy 393 GEDPKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDCRDIIORMHL 452  
Db 297 GPQRDAQAAREFILKMFVDLNP---DSDKIYSHFTCATDTENIRFVFAAVKDTILQLNL 353  
Qy 453 KQYELL 458  
Db 354 KEYNLV 359

## RESULT 15

US-11-060-023-17  
; Sequence 17, Application US/11060023  
; Publication No. US2005025531A1  
; GENERAL INFORMATION:  
; APPLICANT: Aventis Pharma Deutschland GmbH  
; TITLE OF INVENTION: Process for identifying modulators of G protein coupled  
; TITLE OF INVENTION: receptors  
; FILE REFERENCE: AVE D-2000/A033 englisch  
; CURRENT APPLICATION NUMBER: US/11/060,023  
; CURRENT FILING DATE: 2005-02-17  
; PRIOR APPLICATION NUMBER: US/09/899,295  
; PRIOR FILING DATE: 2003-07-06  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-11-060-023-17

Query Match 29.9%; Score 717.5; DB 7; Length 360;  
Best Local Similarity 40.2%; Pred. No. 5e-46;  
Matches 148; Conservative 72; Mismatches 129; Indels 19; Gaps 4;  
Qy 91 EREAVKEARKVSRGIDRMLRDQKRDLOQTHRLLLGAGESGKSTIVKQMRILHVGNFPE 150  
Db 12 QRNEEKEARRINDEIERHVRDRKEDARRELKLLLLGTGESGKSTFIKQMRIHSGYSDE 71  
Qy 151 EKQKILDIRKVKDAIVTVISAMSTIIPVPLANPENQPSDYIKSIAPITDFEYSQEF 210  
Db 72 DKRGFTKLVYQNIPTAMQAMIRAMDTL--KIPYKYEHNKAHAQLVREVDVEKVSFAFENPY 129  
Qy 211 FDHVKKLWDDEGVKACFERSNEYOLIDCAQYFLERIDSLSVDYPTDQDLRCRLVLTSG 270  
Db 130 VDAIKSLWNDPGIOECYDRREYQLSDSTKYKYLNDLDRVADPAYLPTQDVLVRVPTTG 189  
Qy 271 IFETRFQVDKYNHMFVDCGQDRERKWIQCFNDVTAIYVACSSYNMVIREDNNTNRL 330  
Db 190 IIEYFPDLQSVIFRMVDVGQSRERKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 249  
Qy 331 RESLDLFESIMNNRWLRTISIILFNKQDMLAEKVLGKAGKIEDYPEYANYTVPEDATP 390  
Db 250 EESKALFRITIIYFPWFQNSSVILFNKKDLLEKIM--YSHLVDYFPEY----- 296  
Qy 393 DAGEDPKVTRAKFFIRDLFLRISTATGDKGHYCYPHFTCAVDTENIRRVFNDCRDIIORM 450  
Db 297 -DGQRDAQAAREFILKMFVDLNP---DSDKIYSHFTCATDTENIRFVFAAVKDTILQL 352  
Qy 451 HLKQYELL 458  
Db 353 NLKEYNLV 360

Search completed: March 2, 2006, 19:38:17  
Job time : 16.9223 secs

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March. 2, 2006, 19:29:36 ; Search time 9.63599 Seconds  
(without alignments)  
1413.528 Million cell updates/sec

Title: US-10-618-320A-1\_COPY\_96\_126

Perfect score: 153

Sequence: 1 KEARKVSRGIDRLMDQKRLDQTHLLLG 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_21.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*
- 9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153	100.0	150	9	AEAL7334
2	153	100.0	202	9	AEAL7301
3	153	100.0	458	8	ADG74722 Human G-p
4	153	100.0	458	9	AEAL7292
5	153	100.0	461	4	ABG01236
6	149	97.4	230	9	AEAL7296
7	149	97.4	448	8	ADG74746
8	149	97.4	450	8	ADG74747
9	79	51.6	385	8	ADN2325 Bacterial
10	76	49.7	353	4	ABG9072 Human G-p
11	76	49.7	353	5	ABG68587
12	76	49.7	353	5	ABG68585
13	76	49.7	353	5	ABG68599
14	76	49.7	353	5	ABG09281
15	76	49.7	353	7	ABR82631 C. elegans
16	76	49.7	353	7	ADC09616 Human G-p
17	76	49.7	353	8	ADS73825
18	76	49.7	353	8	ADU60735
19	76	49.7	353	9	ADY64330 Mouse G-p
20	76	49.7	359	2	AAAY49125
21	76	49.7	359	5	ABG68598
22	76	49.7	359	5	ABG68603
23	76	49.7	359	5	ABG68586
24	76	49.7	359	5	ABG68584

25	76	49.7	359	5	ABG68593
26	76	49.7	359	5	ABG68610
27	76	49.7	359	7	ADP70780
28	76	49.7	359	8	ADN06136
29	76	49.7	359	9	ADN07219
30	76	49.7	359	9	ABE71353
31	76	49.7	360	8	ADN06150
32	76	49.7	360	8	ADN06148
33	76	49.7	499	9	ADY14842
34	76	49.7	1276	2	AAAY49127
35	76	49.7	1276	5	AAO15093
36	76	49.7	1303	2	AAAY49132
37	76	49.7	1303	5	AAO15098
38	76	49.7	1323	2	AAAY49133
39	76	49.7	1323	5	AAO15099
40	76	49.7	1394	2	AAAY49129
41	76	49.7	1394	5	AAO15095
42	76	49.7	1397	2	AAAY49134
43	76	49.7	1397	5	AAO15100
44	76	49.7	1402	5	AAO15105
45	76	49.7	1418	2	AAAY49131

## ALIGNMENTS

RESULT 1  
AEAL7334  
ID AEAL7334 standard; protein; 150 AA.  
XX  
AC AEAL7334;  
XX  
DT 28-JUL-2005 (first entry)  
XX  
DE N-terminal human XLG(olf) protein fragment Seq 44.  
XX  
KW XLG(olf); G protein coupled receptor; schizophrenia;  
KW psychiatric disorder; neuroleptic; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO2005047318-A1.  
XX  
PD 26-MAY-2005.  
XX  
PF 11-NOV-2004; 2004WO-GB004749.  
XX  
PR 11-NOV-2003; 2003US-0519190P.  
PR 03-SEP-2004; 2004US-0607010P.  
XX  
(ASTR ) ASTRAZENECA AB.  
(ASTR ) ASTRAZENECA UK LTD.  
XX  
PI Bostwick RJ, Corradi J, Defay T, Furlong S, Hirata LT, Ravyn V;  
PI Robbins A;  
XX  
XX WPI; 2005-386333/39.  
XX  
PT New isolated nucleic acid molecule encoding Golf G proteins, useful for  
PT identifying modulators of G protein coupled receptor activity, or for  
PT diagnosing or treating schizophrenia and other psychiatric disorders.  
XX  
PS Disclosure; SEQ ID NO 44; 235pp; English.  
XX  
CC This invention relates to a novel transcriptional variant of the human  
CC GNAL gene that encodes a novel splice variant of the G protein alpha  
CC subunit protein Golf, referred to herein as XLGolf, as well as methods  
CC for identifying modulators of G protein coupled receptors (GPCRs). Note  
CC that the N-terminus of the XLGolf protein is altered compared to Golf  
CC with a different exon 1. Specifically, it refers to contacting the GPCR  
CC with a test compound, and determining GPCR activity, where a change in  
CC activity indicates that the compound is a modulator thereof. The present  
CC invention describes the GPCR as a Gs coupled GPCR that is selected from

Abg68593 Mouse G p  
Abg68610 Human G p  
Adp70780 Minicell  
Adn06136 Human Gq  
Adx07219 Cyclin-de  
Aeb71353 Human mod  
Adn06150 Rat Gii-H  
Adn06148 Rat Gii-H  
Ady14842 PRO polyp  
Aay49127 phCar/hmg  
Aao15093 Human phC  
Aay49132 GABA-BR2\*  
Aao15098 Human GAB  
Aay49133 GABA-BR1a  
Aao15099 Human GAB  
Aay49129 pmGluR2/C  
Aao15095 Human pmG  
Aay49134 pmGluR2/C  
Aao15100 Human pmG  
Aao15105 Human ph2  
Aay49131 mGluR8/Ca

CC dopamine receptor D1, adenosine A2a receptor, and adrenergic beta-2  
 CC receptor. Accordingly, the composition and methods are useful for  
 CC identifying modulators of GPCR activity, as well as for diagnosing or  
 CC treating schizophrenia and other psychiatric disorders. Furthermore, the  
 CC pharmaceutical compositions derived thereof exhibit neuroleptic activity  
 CC and can be used for gene therapy purposes. This polypeptide sequence is  
 CC an N-terminal fragment of the human XLG(olf) protein of the invention.

XX  
 SQ Sequence 150 AA;  
 Query Match 100.0%; Score 153; DB 9; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-13;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDKQKRDLOQTHRLLLG 31  
 |||||  
 DB 96 KEARKVSRGIDRMRLDKQKRDLOQTHRLLLG 126  
 |||||

RESULT 2  
 ID AEA17301 standard; protein; 202 AA.  
 XX AEA17301;  
 AC  
 XX  
 XX 28-JUL-2005 (first entry)  
 DT  
 XX Human XLGalphas protein fragment (encoded by alternate transcript) Seq1.  
 DE  
 XX XLGalphas; G protein coupled receptor; schizophrenia;  
 KW psychiatric disorder; neuroleptic; gene therapy.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO2005047318-A1.  
 PN  
 XX 26-MAY-2005.  
 PD  
 XX 11-NOV-2004; 2004WO-GB004749.  
 PF  
 XX 11-NOV-2003; 2003US-0519190P.  
 PR 03-SEP-2004; 2004US-0607010P.  
 XX (ASTR ) ASTRAZENECA AB.  
 PA (ASTR ) ASTRAZENECA UK LTD.  
 XX  
 XX Bostwick RJ, Corradi J, Defay T, Furlong S, Hirata LT, Ravyn V;  
 PI Robbins A;  
 PT WPI; 2005-386333/39.  
 PS  
 XX New isolated nucleic acid molecule encoding Golf G proteins, useful for  
 PT identifying modulators of G protein coupled receptor activity, or for  
 PT diagnosing or treating schizophrenia and other psychiatric disorders.  
 PS Disclosure; SEQ ID NO 11; 235pp; English.

XX  
 CC This invention relates to a novel transcriptional variant of the human  
 CC GNAL gene that encodes a novel splice variant of the G protein alpha  
 CC subunit protein Golf, referred to herein as XLGolf, as well as methods  
 CC for identifying modulators of G protein coupled receptors (GPCRs). Note  
 CC that the N-terminus of the XLGolf protein is altered compared to Golf  
 CC with a different exon 1. Specifically, it refers to contacting the GPCR  
 CC with a test compound, and determining GPCR activity, where a change in  
 CC activity indicates that the compound is a modulator thereof. The present  
 CC invention describes the GPCR as a Gs coupled GPCR that is selected from  
 CC dopamine receptor D1, adenosine A2a receptor, and adrenergic beta-2  
 CC receptor. Accordingly, the composition and methods are useful for  
 CC identifying modulators of GPCR activity, as well as for diagnosing or  
 CC treating schizophrenia and other psychiatric disorders. Furthermore, the  
 CC pharmaceutical compositions derived thereof exhibit neuroleptic activity  
 CC and can be used for gene therapy purposes. This polypeptide sequence is  
 CC an N-terminal fragment of the human XLGalphas protein of the invention.

XX  
 SQ Sequence 202 AA;  
 Query Match 100.0%; Score 153; DB 9; Length 202;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-13;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDKQKRDLOQTHRLLLG 31  
 |||||  
 DB 96 KEARKVSRGIDRMRLDKQKRDLOQTHRLLLG 126  
 |||||

RESULT 3  
 ID ADG74722 standard; protein; 458 AA.  
 XX ADG74722;  
 AC  
 XX 22-APR-2004 (first entry)  
 DT  
 XX Human G-protein Gml amino acid sequence.  
 DE  
 XX G protein; Gml; G protein-coupled receptor mediated signal transduction;  
 KW GTP binding site; GTPase site; G protein alpha subunit;  
 KW signal transduction; G-protein-coupled receptor.  
 XX  
 XX Homo sapiens.  
 OS  
 XX EP1382613-A1.  
 PN  
 XX 21-JAN-2004.  
 PD  
 XX 09-JUL-2003; 2003EP-00015519.  
 PF  
 XX 16-JUL-2002; 2002JP-00206841.  
 PR 19-DEC-2002; 2002JP-00367778.  
 PR 31-MAR-2003; 2003JP-00095955.  
 XX (SUMO ) SUMITOMO CHEM CO LTD.  
 PA  
 XX Takahashi Y, Matsumoto Y, Oeda K;  
 PI WPI; 2004-111483/12.  
 XX N-PSDB; ADG74723.  
 DR  
 XX  
 XX New protein useful as a therapeutic or prophylactic agent against a  
 PT disease caused by an abnormality in a G-protein coupled receptor mediated  
 PT signal transduction.  
 PT  
 XX Claim 1; SEQ ID NO 1; 85pp; English.

XX  
 CC This invention relates to a novel G protein (Gml). The protein is  
 CC involved in a G protein-coupled receptor mediated signal transduction.  
 CC The protein of the invention has a sequence with a high homology with a  
 CC GTP binding site and a GTPase site conserved among G protein alpha  
 CC subunits. The protein, the DNA sequence which encodes it and an antibody  
 CC specifically recognizing the protein of the invention may be useful as a  
 CC therapeutic or prophylactic agent against a disease caused by an  
 CC abnormality in a G-protein coupled receptor mediated signal transduction.  
 CC The invention may also be useful for screening for a substance capable of  
 CC regulating a signal transduction mediated by a G-protein-coupled receptor  
 CC and a protein.

XX  
 SQ Sequence 458 AA;  
 Query Match 100.0%; Score 153; DB 8; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDKQKRDLOQTHRLLLG 31  
 |||||  
 DB 96 KEARKVSRGIDRMRLDKQKRDLOQTHRLLLG 126  
 |||||

RESULT 4  
 AEA17292  
 ID AEA17292 standard; protein; 458 AA.  
 XX  
 AC AEA17292;  
 DT 28-JUL-2005 (first entry)  
 XX  
 DE Human XLGolf protein Seq 2.  
 XX  
 KW XLGolf; G protein coupled receptor; schizophrenia; psychiatric disorder;  
 KW neuroleptic; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2005047318-A1.  
 XX  
 PD 26-MAY-2005.  
 XX  
 PF 11-NOV-2004; 2004WO-GB004749.  
 XX  
 PR 11-NOV-2003; 2003US-0519190P.  
 PR 03-SEP-2004; 2004US-0607010P.  
 XX  
 PA (ASTR ) ASTRAZENECA AB.  
 PA (ASTR ) ASTRAZENECA UK LTD.  
 XX  
 PI Bostwick RJ, Corradi J, Defay T, Furlong S, Hirata LT, Raym V;  
 PI Robbins A;  
 XX  
 DR WPI; 2005-386333/39.  
 DR N-PSDB; AEA17291, AEA17302.  
 XX  
 PT New isolated nucleic acid molecule encoding Golf G proteins, useful for  
 PT identifying modulators of G protein coupled receptor activity, or for  
 PT diagnosing or treating schizophrenia and other psychiatric disorders.  
 XX  
 PS Claim 8; SEQ ID NO 2; 235pp; English.  
 XX  
 CC This invention relates to a novel transcriptional variant of the human  
 CC GNAL gene that encodes a novel splice variant of the G protein alpha  
 CC subunit protein Golf, referred to herein as XLGolf, as well as methods  
 CC for identifying modulators of G protein coupled receptors (GPCRs). Note  
 CC that the N-terminus of the XLGolf protein is altered compared to Golf  
 CC with a different exon 1. Specifically, it refers to contacting the GPCR  
 CC with a test compound, and determining GPCR activity, where a change in  
 CC activity indicates that the compound is a modulator thereof. The present  
 CC invention describes the GPCR as a Gs coupled GPCR that is selected from  
 CC dopamine receptor D1, adenosine A2a receptor, and adrenergic beta-2  
 CC receptor. Accordingly, the composition and methods are useful for  
 CC identifying modulators of GPCR activity, as well as for diagnosing or  
 CC treating schizophrenia and other psychiatric disorders. Furthermore, the  
 CC pharmaceutical compositions derived thereof exhibit neuroleptic activity  
 CC and can be used for gene therapy purposes. This polypeptide sequence is  
 CC the human XLGolf protein of the invention.  
 XX  
 SQ Sequence 458 AA;  
 Query Match 100.0%; Score 153; DB 9; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KEARKVSRGIDRMRLDKRDLQOQTHRLLLLG 31  
 DB 96 KEARKVSRGIDRMRLDKRDLQOQTHRLLLLG 126

RESULT 5  
 ABG01236  
 ID ABG01236 standard; protein; 461 AA.  
 XX  
 AC ABG01236;  
 DT 28-JUL-2005 (first entry)

XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #1227.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS65423.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 31595; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 461 AA;  
 Query Match 100.0%; Score 153; DB 4; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KEARKVSRGIDRMRLDKRDLQOQTHRLLLLG 31  
 DB 96 KEARKVSRGIDRMRLDKRDLQOQTHRLLLLG 126

RESULT 6  
 AEA17296  
 ID AEA17296 standard; protein; 230 AA.  
 XX  
 AC AEA17296;  
 XX  
 DT 28-JUL-2005 (first entry)

XX Murine XLGolf protein Seq 6.  
DE  
XX  
XX XLGolf; G protein coupled receptor; schizophrenia; psychiatric disorder;  
KW neuroleptic; gene therapy.  
XX  
XX Mus musculus.  
OS  
XX W02005047318-A1.  
PN  
XX  
XX 26-MAY-2005.  
PD  
XX  
XX 11-NOV-2004; 2004WO-GB004749.  
PF  
XX  
XX 11-NOV-2003; 2003US-0519190P.  
PR  
XX 03-SEP-2004; 2004US-0607010P.  
XX  
XX (ASTR ) ASTRAZENECA AB.  
PA (ASTR ) ASTRAZENECA UK LTD.  
PA  
XX Bostwick RJ, Corradi J, Defay T, Furlong S, Hirata LT, Ravyn V;  
PI Robbins A;  
PI  
XX WPI; 2005-386333/39.  
DR N-PSDB; AEA17295.  
XX  
XX New isolated nucleic acid molecule encoding Golf G proteins, useful for  
PT identifying modulators of G protein coupled receptor activity, or for  
PT diagnosing or treating schizophrenia and other psychiatric disorders.  
XX  
XX Example 1; SEQ ID NO 6; 235pp; English.  
PS  
XX This invention relates to a novel transcriptional variant of the human  
CC GNAI1 gene that encodes a novel splice variant of the G protein alpha  
CC subunit protein Golf, referred to herein as XLGolf, as well as methods  
CC for identifying modulators of G protein coupled receptors (GPCRs). Note  
CC that the N-terminus of the XLGolf protein is altered compared to Golf  
CC with a different exon 1. Specifically, it refers to contacting the GPCR  
CC with a test compound, and determining GPCR activity, where a change in  
CC activity indicates that the compound is a modulator thereof. The present  
CC invention describes the GPCR as a Gs coupled GPCR that is selected from  
CC dopamine receptor D1, adenosine A2a receptor, and adrenergic beta-2  
CC receptor. Accordingly, the composition and methods are useful for  
CC identifying modulators of GPCR activity, as well as for diagnosing or  
CC treating schizophrenia and other psychiatric disorders. Furthermore, the  
CC pharmaceutical compositions derived thereof exhibit neuroleptic activity  
CC and can be used for gene therapy purposes. This polypeptide sequence is  
CC the murine XLGolf protein of the invention.  
XX  
XX Sequence 230 AA;  
SQ  
Query Match 97.4%; Score 149; DB 9; Length 230;  
Best Local Similarity 96.8%; Pred. No. 2.1e-12;  
Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEARKVSRGIDRLMDKRDQLQOQTHRLLLG 31  
DB 86 KEARKVSRGIDRLMDKRDQLQOQTHRLLLG 116  
XX  
XX  
XX  
XX  
XX 22-APR-2004 (first entry)  
DE  
XX Mouse G-protein Gml amino acid sequence.  
DE  
XX  
XX G protein; Gml; G protein-coupled receptor mediated signal transduction;  
KW GTP binding site; GTPase site; G protein alpha subunit;  
KW signal transduction; G-protein-coupled receptor; mouse; murine.  
XX

XX Mus musculus.  
OS  
XX EP1382613-A1.  
PN  
XX  
XX 21-JAN-2004.  
PD  
XX  
XX 09-JUL-2003; 2003EP-00015519.  
PF  
XX  
XX 16-JUL-2002; 2002JP-00206841.  
PR 19-DEC-2002; 2002JP-00367778.  
PR 31-MAR-2003; 2003JP-00095955.  
XX  
XX (SUMO ) SUMITOMO CHEM CO LTD.  
PA  
XX Takahashi Y, Matsumoto Y, Oeda K;  
PI WPI; 2004-111483/12.  
DR N-PSDB; ADG74748.  
XX  
XX New protein useful as a therapeutic or prophylactic agent against a  
PT disease caused by an abnormality in a G-protein coupled receptor mediated  
PT signal transduction.  
XX  
XX Claim 1; SEQ ID NO 25; 85pp; English.  
PS  
XX This invention relates to a novel G protein (Gml). The protein is  
CC involved in a G protein-coupled receptor mediated signal transduction.  
CC The protein of the invention has a sequence with a high homology with a  
CC GTP binding site and a GTPase site conserved among G protein alpha  
CC subunits. The protein, the DNA sequence which encodes it and an antibody  
CC specifically recognizing the protein of the invention may be useful as a  
CC therapeutic or prophylactic agent against a disease caused by an  
CC abnormality in a G-protein coupled receptor mediated signal transduction.  
CC The invention may also be useful for screening for a substance capable of  
CC regulating a signal transduction mediated by a G-protein-coupled receptor  
CC and a protein. The present sequence is that of the mouse Gml protein  
CC which is related to the human Gml protein of the invention.  
XX  
XX Sequence 448 AA;  
SQ  
Query Match 97.4%; Score 149; DB 8; Length 448;  
Best Local Similarity 96.8%; Pred. No. 4.2e-12;  
Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEARKVSRGIDRLMDKRDQLQOQTHRLLLG 31  
DB 86 KEARKVSRGIDRLMDKRDQLQOQTHRLLLG 116  
XX  
XX  
XX  
XX 22-APR-2004 (first entry)  
DT  
XX Rat G-protein Gml amino acid sequence.  
DE  
XX  
XX G protein; Gml; G protein-coupled receptor mediated signal transduction;  
KW GTP binding site; GTPase site; G protein alpha subunit;  
KW signal transduction; G-protein-coupled receptor; rat.  
XX  
XX Rattus norvegicus.  
OS  
XX EP1382613-A1.  
PN  
XX  
XX 21-JAN-2004.  
PD  
XX  
XX 09-JUL-2003; 2003EP-00015519.  
PF  
XX  
XX 16-JUL-2002; 2002JP-00206841.  
PR



```
PR 19-DEC-2002; 2002JP-00367778.
PR 31-MAR-2003; 2003JP-00095955.
XX (SUMO) SUMITOMO CHEM CO LTD.
XX Takahashi Y, Matsumoto Y, Oeda K;
XX WPI; 2004-111483/12.
DR N-PSDB; ADG74749.
XX
PT New protein useful as a therapeutic or prophylactic agent against a
PT disease caused by an abnormality in a G-protein coupled receptor mediated
PT signal transduction.
XX
PS Claim 1; SEQ ID NO 26; 85pp; English.
XX
CC This invention relates to a novel G protein (Gml). The protein is
CC involved in a G protein-coupled receptor mediated signal transduction.
CC The protein of the invention has a sequence with a high homology with a
CC GTP binding site and a GTPase site conserved among G protein alpha
CC subunits. The protein, the DNA sequence which encodes it and an antibody
CC specifically recognising the protein of the invention may be useful as a
CC therapeutic or prophylactic agent against a disease caused by an
CC abnormality in a G-protein coupled receptor mediated signal transduction.
CC The invention may also be useful for screening for a substance capable of
CC regulating a signal transduction mediated by a G-protein-coupled receptor
CC and a protein. The present sequence is that of the rat Gml protein which
CC is related to the human Gml protein of the invention.
XX
SQ Sequence 450 AA;
Query Match 97.4%; Score 149; DB 8; Length 450;
Best Local Similarity 96.8%; Pred. No. 4.2e-12;
Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KEARKVSRGIDRLMDKRDLDQOQTHRLLLIG 31
DB 88 KEARKVSRGIDRLMDKRDLDQOQTHRLLLIG 118
RESULT 9
ADN22325
ID ADN22325 standard; protein; 385 AA.
XX
AC ADN22325;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #4978.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
CAOY/) CAO Y.
(HINK/) HINKLE G J.
(SLAT/) SLATER S C.
(CHEN/) CHEN X.
(GOLD/) GOLDMAN B S.
XX
```

```
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 4978; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 385 AA;
Query Match 51.6%; Score 79; DB 8; Length 385;
Best Local Similarity 51.6%; Pred. No. 0.018;
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 KEARKVSRGIDRLMDKRDLDQOQTHRLLLIG 31
DB 16 REARKVKKQIEQLAKDKQVWRATHRLLLIG 46
RESULT 10
AAB99072
ID AAB99072 standard; protein; 353 AA.
XX
AC AAB99072;
XX
DT 23-AUG-2001 (first entry)
XX
DE Human G-protein alpha subunit q.
XX
KW G-protein coupled receptor; GPCR; GNRH receptor; disease treatment;
KW gonadotrophin releasing; hormone receptor; hormone dependent cancer;
KW human; catfish; goldfish; cow; sheep; horse; fruitfly; pig; rat; mouse;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO200136446-A2.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-GB004385.
XX
PR 17-NOV-1999; 99GB-00027215.
XX
(UYBR-) UNIV BRISTOL.
PA
```



ABG68599  
ID ABG68599 standard; protein; 353 AA.  
XX  
AC ABG68599;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Human G protein alpha sub-unit q family variant #1.  
XX  
KW G q protein; sensory signaling; chemoreceptor; tastant; olfactant;  
KW pheromone; G protein alpha sub-unit; q family; G alpha q.  
XX  
OS Homo sapiens.  
XX  
PN WO200236622-A2.  
XX  
PD 10-MAY-2002.  
XX  
PF 24-OCT-2001; 2001WO-US032619.  
XX  
PR 30-OCT-2000; 2000US-0243770P.  
XX  
PA (SENO-) SENOMYX INC.  
XX  
PI Yao Y, Xu H;  
XX  
DR WPI; 2002-519234/55.  
XX  
PT New G-alpha-q protein variants, useful for analyzing and discovering  
PT agonists or antagonists of chemoreceptors, such as G protein coupled  
PT receptors involved in sensing of tastants, olfactants or pheromones.  
XX  
PS Claim 13; Page 28; 32pp; English.  
XX  
CC The invention describes an isolated variant of a G q protein, which  
CC exhibits increased promiscuity relative to the corresponding G q protein.  
CC The variant is used to identify a compound that modulates sensory  
CC signaling in sensory cells and to identify a compound that interacts with  
CC the G q variant protein. The G q protein variant is useful for analysing  
CC and discovering agonists or antagonists of chemoreceptors, such as G  
CC protein coupled receptors involved in sensing of tastants, olfactants or  
CC pheromones. This is the amino acid sequence of a G protein alpha sub-unit  
CC q family (G alpha q) variant that can functionally couple to sensory cell  
CC receptors such as taste GPCR's (G protein-coupled receptors) and  
CC olfactory GPCR's in an overly promiscuous manner  
XX  
SQ Sequence 353 AA;  
Query Match 49.7%; Score 76; DB 5; Length 353;  
Best Local Similarity 51.6%; Pred. No. 0.043;  
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;  
QY 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLIG 31  
Db 10 KEARRINDEIERQLRDKRDARRELKLLIG 40  
RESULT 14  
ABR09281  
ID ABR09281 standard; protein; 353 AA.  
XX  
AC ABR09281;  
XX  
DT 10-JUL-2002 (first entry)  
XX  
DE G protein-coupled receptor (GPCR) >G(Q) SEQ ID NO:27.  
XX  
KW Target activated nucleic acid biosensor; signalling moiety; GPCR;  
KW nucleic acid sensor; detection; engineering; drug optimisation;  
KW G protein-coupled receptor.  
XX  
OS Homo sapiens.  
XX

PN WO200222882-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 13-SEP-2001; 2001WO-US028835.  
XX  
PR 13-SEP-2000; 2000US-0232454P.  
XX  
PA (ARCH-) ARCHEMIX CORP.  
XX  
PI Stanton M, Epstein D, Hamaguchi N;  
XX  
DR WPI; 2002-393977/42.  
XX  
PT Nucleic acid sensor for detecting target molecule, comprises target  
PT molecule activation site and optical signaling unit that changes its  
PT optical properties upon allosteric modulation sensor after recognition of  
PT target.  
XX  
PS Example 12; Page 90; 144pp; English.  
XX  
CC The present invention describes a nucleic acid sensor molecule (I)  
CC comprising a target molecule activation site comprising a structure that  
CC recognises a target molecule and an optical signalling unit including at  
CC least one nucleotide coupled to a signalling moiety that changes its  
CC optical properties upon allosteric modulation of (I) following  
CC recognition of the target molecule. (I) is useful for detecting a target  
CC molecule associated with a pathological condition or genetic alteration.  
CC (I) is useful for identifying a drug compound, by identifying a nucleic  
CC acid biosensor-based molecule profile of target molecules associated with  
CC a disease trait in a patient, administering a candidate compound to the  
CC patient, and monitoring changes in the profile. Alternatively, the method  
CC involves identifying a number of pathway target molecules, administering  
CC a candidate compound to a patient having a disease trait, and monitoring  
CC changes in the structure, level or activity of two or more of the pathway  
CC target molecules using (I). The profile of target molecules or the  
CC changes in the structure is compared to the profile of a reference  
CC healthy or diseased population. (I) is useful in multiple assays, for the  
CC detection of target molecule. (I) is also useful in diagnostic  
CC applications and drug optimisation. The present sequence represents a G  
CC protein-coupled receptor, which is used in an example from the present  
XX  
SQ Sequence 353 AA;  
Query Match 49.7%; Score 76; DB 5; Length 353;  
Best Local Similarity 51.6%; Pred. No. 0.043;  
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;  
QY 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLIG 31  
Db 10 KEARRINDEIERQLRDKRDARRELKLLIG 40  
RESULT 15  
ABR82631  
ID ABR82631 standard; protein; 353 AA.  
XX  
AC ABR82631;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE C. elegans EGL-30 protein related fragment G(q).  
XX  
KW RGS; G-protein; regulator of G-protein signaling; Galphag; uropathic;  
KW antidepressant; tranquillizer; antiarrhythmic; relaxant; EGL-30;  
KW nematode.  
XX  
OS Caenorhabditis elegans.  
XX  
PN WO2003063784-A2.  
XX  
PD 07-AUG-2003.



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein.: protein search, using sw model

Run on: March 2, 2006, 19:30:27 ; Search time 1.83845 Seconds  
(without alignments)  
1622.414 Million cell updates/sec

Title: US-10-618-320A-1\_COPY\_96\_126  
Perfect score: 153  
Sequence: 1 KEARKVSRGIDRMRLRDKEDLQOTHRLLLLG 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	51.6	375	2 T37245	GTP-binding regula
2	79	51.6	385	2 D87723	protein R06A10.2
3	76	49.7	359	2 S71963	GTP-binding protei
4	76	49.7	359	2 S45899	GTP-binding regula
5	72	47.1	846	2 S52418	GTP-binding regula
6	71	46.4	359	1 RGM5Q	GTP-binding protei
7	70	45.8	355	2 A40891	GTP-binding protei
8	68	44.4	353	2 B40891	GTP-binding protei
9	68	44.4	353	2 S34347	GTP-binding regula
10	68	44.4	354	2 S33309	GTP-binding regula
11	68	44.4	359	1 RGHUG	GTP-binding regula
12	68	44.4	359	1 RGM511	GTP-binding regula
13	68	44.4	359	2 S30359	GTP-binding regula
14	68	44.4	360	2 JN0115	GTP-binding regula
15	67	43.8	381	2 I53271	G-protein alpha-o
16	65	42.5	355	2 A41534	GTP-binding protei
17	64	41.8	32	2 S36809	GTP-binding regula
18	64	41.8	91	2 A46685	GTP-binding regula
19	64	41.8	359	2 S45700	G-alpha-11 protein
20	64	41.8	377	1 RGM5A1	GTP-binding regula
21	64	41.8	379	2 A42964	guanine nucleotide
22	64	41.8	380	1 RGHUAI	GTP-binding regula
23	64	41.8	394	1 RGHVAE	GTP-binding regula
24	64	41.8	394	1 RGM5A2	GTP-binding regula
25	64	41.8	394	1 RGHYA2	GTP-binding regula
26	64	41.8	394	1 RGR7A2	GTP-binding regula
27	64	41.8	394	2 S33458	GTP-binding regula
28	64	41.8	395	1 RGHUAI	GTP-binding regula
29	64	41.8	397	1 RGPCA2	GTP-binding regula

30	63	41.2	379	1 RGXLA	GTP-binding regula
31	63	41.2	379	2 A41095	GTP-binding regula
32	63	41.2	381	2 A48071	guanine nucleotide
33	63	41.2	394	1 RGHUGA	GTP-binding regula
34	62	40.5	355	1 RGHUGX	GTP-binding regula
35	62	40.5	355	1 RGR7GX	GTP-binding regula
36	61	39.9	382	1 RGF7AS	GTP-binding regula
37	61	39.9	385	1 RGF7AL	GTP-binding regula
38	58	37.9	374	2 B41534	GTP-binding protei
39	58	37.9	645	2 S14992	dnak-type molecula
40	58	37.9	646	2 T46650	heat shock protei
41	58	37.9	648	2 S53498	dnak-type molecula
42	58	37.9	785	2 AB1582	Muts protein (Muts
43	57	37.3	652	2 A53163	dnak-type molecula
44	57	37.3	785	2 AH1228	Muts protein (Muts
45	56	36.6	355	2 T15288	hypothetical prote

ALIGNMENTS

RESULT 1

T37245

GTP-binding regulatory protein Gs alpha-S chain (adenylate cyclase-stimulating) - Caenorh  
N;Alternate names: G protein a(s) subunit

C;Species: Caenorhabditis elegans

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C;Accession: T37243; T37245

R;Park, J.; Ohshima, S.; Tani, T.; Ohshima, Y.

Gene 194, 183-190, 1997

A;Title: Structure and expression of the gsa-1 gene encoding a G protein alpha(s) subunit

A;Reference number: Z21641; MUID:97417487; PMID:9272860

A;Accession: T37243

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-375 <PAR>

A;Cross-references: UNIPROT:O18678; UNIPARC:UPI0000082C78; EMBL:AB003487; PIDN:BAA22402.1

C;Experimental source: strain N2; mixed developmental stages

A;Accession: T37245

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-375 <PA2>

A;Cross-references: UNIPARC:UPI0000082C78; EMBL:AB003487; PIDN:BAA22402.1

C;Experimental source: strain N2; mixed developmental stages

C;Genetics: gsa-1

A;Map position: I

A;Introns: 31/3; 85/3; 125/3; 158/2; 201/2; 246/2; 327/3

C;Superfamily: GTP-binding regulatory protein Gs alpha chain

C;Keywords: GTP binding; signal transduction

Query Match 51.6%; Score 79; DB 2; Length 375;  
Best Local Similarity 51.6%; Pred. No. 0.0028; 8; Indels 0; Gaps 0;  
Matches 16; Conservative 7; Mismatches 8;

OY 1 KEARKVSRGIDRMRLRDKEDLQOTHRLLLLG 31

Db 13 REARKVVKQIEQLAKDKQVMRATHRLLLLG 43

RESULT 2

D87723

protein R06A10.2 [imported] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004

C;Accession: D87723

R;anonymous, The C. elegans Sequencing Consortium.

A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A;Reference number: A75000; MUID:99069613; PMID:9851916

A;Note: see websites genome.wustl.edu/gsc/Celegans/ and www.sanger.ac.uk/projects/c\_eleg

A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A;Accession: D87723

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-385 <STO>  
A;Cross-references: UNIPROT:O18678; UNIPARC:UPI0000177887; GB:chr\_I; PIDN:AAB96733.1; PID:  
C;Genetics:  
A;Gene: R06A10.2  
A;Map position: 1  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain

Query Match 51.6%; Score 79; DB 2; Length 385;  
Best Local Similarity 51.6%; Pred. No. 0.0029; 8; Indels 0; Gaps 0;  
Matches 16; Conservative

QY 1 KEARKVSRGIDRMRLDQKRDLDQTHRLLLG 31  
Db 16 KEARKVSKIEEQQLAKDQKQWRATHRLLLG 46

RESULT 3  
S71963  
GTP-binding protein alpha-q - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 02-Feb-2001  
C;Accession: S71963  
R;Johnson, G.J.; Leite, L.A.; Dunlop, P.C.  
Biochem. J. 318, 1023-1031, 1996  
A;Title: Specificity of G-alpha(q) and G-alpha(11) gene expression in platelets and erythrocytes  
A;Reference number: S71963; MUID:96433124; PMID:8836152  
A;Accession: S71963  
A;Molecule type: mRNA  
A;Residues: 1-359 <JOH>  
A;Cross-references: UNIPARC:UPI000016AC08; EMBL:L76256; NID:g1478071; PIDN:AAB39498.1; PID:  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
C;Keywords: GTP binding; nucleotide binding; P-loop; signal transduction  
F;46-53/Region: nucleotide-binding motif A (P-loop)  
F;274-277/Region: GTP-binding NKXD motif

Query Match 49.7%; Score 76; DB 2; Length 359;  
Best Local Similarity 51.6%; Pred. No. 0.0069; 8; Indels 0; Gaps 0;  
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDQKRDLDQTHRLLLG 31  
Db 16 KEARRINDEIERQRDKRDARRELKLLLG 46

RESULT 4  
S45699  
GTP-binding regulatory protein alpha chain q - African clawed frog  
N;Alternate names: G-alpha-q protein  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 10-Dec-1994 #sequence\_revision 10-Nov-1995 #text\_change 02-Feb-2001  
C;Accession: S45699  
R;Shapira, H.; Way, J.; Lipinsky, D.; Oron, Y.; Battey, J.F.  
FEBS Lett. 348, 89-92, 1994  
A;Title: Neuromedin B receptor, expressed in Xenopus laevis oocytes, selectively couples to G-protein  
A;Reference number: S45699; MUID:94289861; PMID:8026589  
A;Accession: S45699  
A;Molecule type: mRNA  
A;Residues: 1-359 <SHA>  
A;Cross-references: UNIPARC:UPI0000171485; GB:U10502; NID:g505689; PIDN:AAA52189.1; PID:  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
C;Keywords: GTP binding; nucleotide binding; P-loop; signal transduction  
F;46-53/Region: nucleotide-binding motif A (P-loop)  
F;274-277/Region: GTP-binding NKXD motif  
F;52/Binding site: GTP (Lys) #status predicted  
F;183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 49.7%; Score 76; DB 2; Length 359;  
Best Local Similarity 51.6%; Pred. No. 0.0069; 8; Indels 0; Gaps 0;  
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDQKRDLDQTHRLLLG 31

Db 16 KEARRINDEIERQRDKRDARRELKLLLG 46

RESULT 5  
S52418  
GTP-binding regulatory protein Gs alpha-XL chain - rat  
N;Alternate names: G protein XL-alpha-s  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 14-Jul-1995 #sequence\_revision 10-Nov-1995 #text\_change 02-Feb-2001  
C;Accession: S52418  
R;Kehlenbach, R.H.; Matthey, J.; Huttner, W.B.  
Nature 372, 804-809, 1994  
A;Title: XL-alpha-s is a new type of G protein.  
A;Reference number: S52418; MUID:95089824; PMID:7997272  
A;Accession: S52418  
A;Molecule type: mRNA  
A;Residues: 1-846 <KEH>  
A;Cross-references: UNIPARC:UPI000017C91F; EMBL:X84047; NID:g642267; PIDN:CAA58866.1; PID:  
R;Kehlenbach, R.H.; Matthey, J.; Huttner, W.B.  
Nature 375, 253, 1995  
A;Title: Correction: Xlalphas is a new type of G protein.  
A;Reference number: S58911  
A;Contents: annotation; assignment of start\_codon  
A;Note: experimental data from this paper suggest that the translation is initiated at pC  
C;Keywords: GTP binding; nucleotide binding; P-loop; signal transduction  
F;132-846/Product: GTP-binding regulatory protein Gs alpha-XL chain #status experimental  
F;499-506/Region: nucleotide-binding motif A (P-loop)  
F;744-747/Region: GTP-binding NKXD motif

Query Match 47.1%; Score 72; DB 2; Length 846;  
Best Local Similarity 53.6%; Pred. No. 0.063; 8; Indels 0; Gaps 0;  
Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 RKVSRGIDRMRLDQKRDLDQTHRLLLG 31  
Db 472 KRRSLIDKQLEEKDYMCVTHRLLLG 499

RESULT 6  
RGMSQ  
GTP-binding regulatory protein Gq alpha chain - mouse  
N;Alternate names: guanine nucleotide binding protein Gq alpha chain; heterotrimeric G-p  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
C;Accession: A38414  
R;Strathmann, M.; Simon, M.I.  
Proc. Natl. Acad. Sci. U.S.A. 87, 9113-9117, 1990  
A;Title: G protein diversity: a distinct class of alpha subunits is present in vertebrate  
A;Reference number: A38414; MUID:91067657; PMID:2123549  
A;Accession: A38414  
A;Molecule type: mRNA  
A;Residues: 1-359 <STR>  
A;Cross-references: UNIPROT:P21279; UNIPARC:UPI0000161D60; GB:M55412; NID:g193501; PIDN:  
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay  
cins. The beta and gamma chains, required for GTPase activity, appear to be common to all  
rase; it is specific for each type of G protein.  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
C;Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction  
F;46-53/Region: nucleotide-binding motif A (P-loop)  
F;274-277/Region: GTP-binding NKXD motif  
F;52/Binding site: GTP (Lys) #status predicted  
F;183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 46.4%; Score 71; DB 1; Length 359;  
Best Local Similarity 48.4%; Pred. No. 0.034; 8; Indels 0; Gaps 0;  
Matches 15; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDQKRDLDQTHRLLLG 31  
Db 16 KEARRINDEIERHVRDKRDARRELKLLLG 46

A;Cross-references: UNIPROT:P38411; UNIPARC:UPI000012B263; EMBL:Z23106; NID:gsl2629; PID:  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
C;Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction  
F:40-48/Region:nucleotide-binding motif A (P-loop)  
F:150-152/Region: GTP-binding SAK/L motif  
F:268-271/Region: GTP-binding NKXD motif  
F:46/Binding site: GTP (lys) #ssatus predicted  
F:177/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 44.4%; Score 68; DB 2; Length 353;  
Best Local Similarity 41.9%; Pred. No. 0.087;  
Matches 13; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

Oy 1 KEARKVSGIDRMRLDQKRDQQTHRLLLLG 31  
|| :|:: || |:: :||||  
Db 10 KEQRINQEIERQLKRDKRDRRELKLLLG 40

RESULT 10  
S33309  
GTP-binding regulatory protein Gq alpha chain - northern European squid  
C;Species: Loligo forbesi (northern European squid)  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C;Accession: S33309  
F:RYda, N.J.P.; Findlay, J.B.C.; Reid, J.D.  
Biochem. J. 292, 333-341, 1993  
A;Title: The molecular cloning of the squid (Loligo forbesi) visual Gq-alpha subunit and  
A;Reference number: S33309; MUID:93277493; PMID:9503868  
A;Accession: S33309  
A>Status: preliminary  
A:Molecule type: mRNA  
A;Residues: 1-354 (#RVB)  
A;Cross-references: UNIPROT:P38412; UNIPARC:UPI000012B262; EMBL:L10289  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
C;Keywords: GTP binding; nucleotide binding; P-loop  
F:40-47/Region: nucleotide-binding motif A (P-loop)  
F:150-152/Region: GTP-binding SAK/L motif  
F:269-272/Region: GTP-binding NKXD motif

Query Match 44.4%; Score 68; DB 2; Length 354;  
Best Local Similarity 41.9%; Pred. No. 0.088;  
Matches 13; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

Oy 1 KEARKVSGIDRMRLDQKRDQQTHRLLLLG 31  
|| :|:: || |:: :||||  
Db 10 KEQRINQEIERQLKRDKRDRRELKLLLG 40

RESULT 11  
RGHUGY  
GTP-binding regulatory protein Gy alpha chain - human  
N;Alternate names: guanine nucleotide binding protein Gy alpha chain; heterotrimeric G-protein  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
C;Accession: A39394  
F:Jiang, M.; Pandey, S.; Tran, V.T.; Fong, H.K.W.  
Proc. Natl. Acad. Sci. U.S.A. 89, 3907-3911, 1991  
A;Title: Guanine nucleotide-binding regulatory proteins in retinal pigment epithelial cel  
A;Reference number: A39394; MUID:91219481; PMID:1902575  
A;Accession: A39394  
A:Molecule type: mRNA  
A;Residues: 1-359 (#JTA)  
A;Cross-references: UNIPROT:P29592; UNIPARC:UPI0000161B28; GB:M69013; NID:g183690; PIDN:/  
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay  
ains. The beta and gamma chains, required for GTPase activity, appear to be common to all  
rase; it is specific for each type of G protein.

A;Gene: GDB:GNAl1  
A;Cross-references: GDB:132587; OMIM:139313  
A;Map position: 19p13.3-19p13.3  
C;Superfamily: GTP-binding regulatory protein Gs alpha chain  
C;Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction  
F:46-53/Region: nucleotide-binding motif A (P-loop)





19 KERREANKKIEKQLQKERLAYKATHRLLLLG 49

Search completed: March 2, 2006, 19:31:05  
Job time : 1.83845 secs

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:29:50 ; Search time 19.3988 Seconds  
(without alignments)  
1127.462 Million cell updates/sec

Title: US-10-618-320a-1\_copy\_96\_126

Perfect score: 153

Sequence: 1 KEARKVSRGIDRLMDKQRLDQTHRLLLG 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_05\_80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153	100.0	458	2	Q6XU3 HUMAN
2	149	97.4	230	2	Q8BH8 MOUSE
3	149	97.4	448	2	Q6L47 MOUSE
4	115	75.2	462	2	Q6GL1 XENLA
5	95	62.1	351	2	Q8N24 HUMAN
6	79	51.6	374	2	Q6PWS CAENORHABDI
7	79	51.6	375	2	Q4VT24 CAEBR
8	79	51.6	375	2	Q18678 CAEBR
9	78	51.0	485	2	Q4SQ90 TETNG
10	76	49.7	353	1	GNAQ CANFA
11	76	49.7	353	1	GNAQ HUMAN
12	76	49.7	353	1	GNAQ MOUSE
13	76	49.7	353	1	GNAQ RAT
14	76	49.7	359	2	Q6NT27 HUMAN
15	76	49.7	359	2	Q5F3B5 CHICK
16	76	49.7	359	2	Q6L128 XENLA
17	76	49.7	502	2	Q4RT60 TETNG
18	73	47.7	377	2	Q6TP31 PINFU
19	72	47.1	177	2	Q9WUC0 MOUSE
20	72	47.1	353	1	GNAQ HUMAN
21	72	47.1	353	1	GNAQ XENLA
22	72	47.1	353	2	Q91955 LIMPO
23	72	47.1	353	2	Q58FC0 HELAM
24	72	47.1	353	2	Q8T6P8 WAMER
25	72	47.1	353	2	Q9U473 PANAR
26	72	47.1	353	2	Q6ITD0 PENVA
27	72	47.1	353	2	Q76FN3 BOMWO
28	72	47.1	359	2	Q5RKP9 BRARE
29	72	47.1	715	2	Q63803 RAT
30	72	47.1	756	2	Q9ZLN8 MOUSE
31	72	47.1	827	2	Q6R0H4 MOUSE

32	72	47.1	876	2	Q6R0H5 MOUSE
33	72	47.1	1133	2	Q6R0H7 MOUSE
34	71	46.4	353	2	Q9NF20 CALVI
35	71	46.4	353	2	Q9NL92 OCTVU
36	71	46.4	353	2	Q7PHK1 ANOGA
37	71	46.4	353	2	Q7PHK1 ANOGA
38	71	46.4	358	2	Q6W9M4 PENVA
39	71	46.4	367	2	Q4RMQ6 TETNG
40	70	45.8	92	2	Q5JW64 HUMAN
41	70	45.8	185	2	Q5JWE8 HUMAN
42	70	45.8	188	2	Q75684 HUMAN
43	70	45.8	193	2	Q5JWE9 HUMAN
44	70	45.8	355	1	GNA14 BOVIN
45	70	45.8	355	1	GNA14 HUMAN

#### ALIGNMENTS

RESULT 1  
Q86XU3 HUMAN PRELIMINARY; PRT; 458 AA.  
AC Q86XU3;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Guanine nucleotide binding protein (G protein), alpha activating  
DE activity polypeptide, olfactory type, isoform 1.  
GN Names=GNAL;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Testis;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Testis;  
RG NIH MGC Project;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
EMBL: BC050021; AAH50021.1; -, mRNA.  
DR HSSP; Q04896; 1CJU.  
DR SMR; Q86XU3; 117-455.  
DR Ensembl; ENSG00000141404; Homo sapiens.  
DR GO; GO:0005525; F:GTP binding; IEA.  
DR GO; GO:0004871; F:signal transducer activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . . ; IEA.  
DR GO; GO:0007165; P:signal transduction; IEA.  
DR InterPro; IPR01019; Gprotein\_alpha\_bd.  
DR InterPro; IPR000367; Gprotein\_alpha\_s.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Testis;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Testis;  
RG NIH MGC Project;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
EMBL: BC050021; AAH50021.1; -, mRNA.  
DR HSSP; Q04896; 1CJU.  
DR SMR; Q86XU3; 117-455.  
DR Ensembl; ENSG00000141404; Homo sapiens.  
DR GO; GO:0005525; F:GTP binding; IEA.  
DR GO; GO:0004871; F:signal transducer activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . . ; IEA.  
DR GO; GO:0007165; P:signal transduction; IEA.  
DR InterPro; IPR01019; Gprotein\_alpha\_bd.  
DR InterPro; IPR000367; Gprotein\_alpha\_s.



DR GO; GO:0007190; P:adenylylate cyclase activation; IDA.  
 DR GO; GO:0007608; P:perception of smell; IMP.  
 DR InterPro; IPR011025; Gprotein\_insd.  
 DR InterPro; IPR001019; Gprotein\_alpha.  
 DR InterPro; IPR000367; Gprotein\_alpha\_s.  
 DR Pfam; PF00503; G-alpha; 1.  
 DR PRINTS; PR00443; GPROTEINAS.  
 DR PRODOM; PD000281; Gprotein\_alpha; 1.  
 DR SMART; SM00275; G\_alpha; 1.  
 SQ SEQUENCE 230 AA; 25632 MW; 0BE3051094E326C2 CRC64;

Query Match 97.4%; Score 149; DB 2; Length 230;  
 Best Local Similarity 96.8%; Pred. No. 2; 8e-12;  
 Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLRQKRDQLQOQTHRLLLG 31  
 |||||  
 DB 86 KEARKVSRGIDRMRLRQKRDQLQOQTHRLLLG 116

RESULT 3  
 Q66L47 MOUSE  
 ID Q66L47 MOUSE PRELIMINARY; PRT; 448 AA.  
 AC Q66L47;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 23-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 23-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Gnal protein.  
 GN Names=Gnal;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridea; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RG NIH MGC Project;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR ENBL; BC078439; AAH78439.1; -, mRNA.  
 DR SNR; Q66L47; 107-445.  
 DR Ensemble; ENSMUSG00000024524; Mus musculus.  
 DR MGI; MGI:95774; Gnal.  
 DR GO; GO:0007190; P:adenylylate cyclase activation; IDA.  
 DR GO; GO:0007608; P:perception of smell; IMP.  
 DR InterPro; IPR001019; Gprotein\_alpha.  
 DR InterPro; IPR000367; Gprotein\_alpha\_s.  
 DR InterPro; IPR011025; Gprotein\_insd.  
 DR Pfam; PF00503; G-alpha; 1.

DR PRINTS; PR00318; GPROTEINA.  
 DR PRINTS; PR00443; GPROTEINAS.  
 DR PRODOM; PD000281; Gprotein\_alpha; 1.  
 DR SMART; SM00275; G\_alpha; 1.  
 SQ SEQUENCE 448 AA; 51357 MW; AE192038B7E3B40E CRC64;

Query Match 97.4%; Score 149; DB 2; Length 448;  
 Best Local Similarity 96.8%; Pred. No. 5; 6e-12;  
 Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLRQKRDQLQOQTHRLLLG 31  
 |||||  
 DB 86 KEARKVSRGIDRMRLRQKRDQLQOQTHRLLLG 116

RESULT 4  
 Q6GLL1 XENLA  
 ID Q6GLL1 XENLA PRELIMINARY; PRT; 462 AA.  
 AC Q6GLL1;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE LOC443714 protein (Fragment).  
 GN Names=LOC443714;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 Dev. Dyn. 225:384-391 (2002).  
 RL [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RA Klein S.; Gerhard D.S.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR ENBL; BC074466; AAH74466.1; -, mRNA.  
 DR SNR; Q6GLL1; 120-459.  
 DR GO; GO:0005525; F:GTP binding; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin.; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.

```
DR InterPro: IPR011025; G-proteinA insert.
DR InterPro: IPR001019; G-protein_alpha insert.
DR InterPro: IPR000367; G-protein_alpha_s.
DR Pfam: PF00503; G-alpha; 1.
DR PRINTS: PR00318; GPROTEINA.
DR ProDom: PD000281; G-protein_alpha; 1.
DR SMART: SM00275; G_alpha; 1.
FT NON TER
SQ SEQUENCE 462 AA; 53392 MW; 57EB03C65C17DAF4 CRC64;

Query Match 75.28; Score 115; DB 2; Length 462;
Best Local Similarity 71.08; Pred. No. 2.9e-07;
Matches 25; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRLMDKQKRDLOQTHRLLLLG 31
   |||||:|||||:|||||:|||||:
DB 100 KEAKVSKTTDRVLKEQKREYKQTHRLLLLG 130

RESULT 5
Q8N2B4_HUMAN
ID Q8N2B4_HUMAN PRELIMINARY; PRT; 351 AA.
AC Q8N2B4_2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein FLJ33549.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Amalgama; DOI=10.1038/ng1285;
RX PubMed=14702039;
RA Oka T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayaashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwanoagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kuano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakami B.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami Y.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45 (2004).
DR EMBL; AK090868; BAC03535.1; -; mRNA.
DR HSP; P04896; 1GJU.
DR SMR; Q8N2B4; 10-348.
```



DT 13-SEP-2005 (TReMBLrel. 31, Created)  
 DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)  
 DE Gsa-1.  
 OS Caenorhabditis briggsae.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6238;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=AF16;  
 RA Jovelin R., Phillips P.C.;  
 RT "Functional constraint and divergence in the G protein family in  
 RT Caenorhabditis elegans and Caenorhabditis briggsae."  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY634306; AAW02912.1; -; Genomic DNA.  
 SQ SEQUENCE 375 AA; 43979 MW; DE43530DAD61D42A CRC64;  
  
 Query Match 51.6%; Score 79; DB 2; Length 375;  
 Best Local Similarity 51.6%; Pred. No. 0.022;  
 Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;  
  
 Qy 1 KEARKVSRGIDRMLRDKQKRDLOQTHRLLLLG 31  
 Db 13 REARKVKNQIEQLAKDKQVMRATHRLLLLG 43  
  
 RESULT 8  
 O18678 CAEEL  
 ID O18678 CAEEL PRELIMINARY; PRT; 375 AA.  
 AC O18678; 044744;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)  
 DE Heterotrimeric G protein alpha subunit (G protein, subunit alpha  
 DE protein 1) (G protein a(S) subunit).  
 GN Name=gsa-1; ORFNames=R06A10.2;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=N2;  
 RX MEDLINE=37417487; PubMed=9272860; DOI=10.1016/S0378-1119(97)00122-4;  
 RA Park J., Ohshima S., Tani T., Ohshima Y.;  
 RT "Structure and expression of the gsa-1 gene encoding a G protein  
 RT alpha(s) subunit in C. elegans."  
 RL Gene 194:183-190(1997).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RG The C. elegans sequencing consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=N2;  
 RA Cuppen E., Jansen G., Plasterk R.H.A.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB003487; BAA22402.1; -; Genomic DNA.  
 DR EMBL; AF019715; AAB96733.2; -; Genomic DNA.  
 DR EMBL; AY008141; AAC32094.1; -; mRNA.  
 DR EMBL; AB003486; BAA22401.1; -; mRNA.  
 DR PIR; D87723; D87723.  
 DR PIR; T37243; T37245.  
 DR HSSP; P04896; 1CTU.  
 DR SRR; O18678; 31-372.  
 DR Ensembl; R06A10.2; Caenorhabditis elegans.  
 DR WormBase; WBGene00001745; gsa-1.  
 DR WormPep; R06A10.2; CE21115.

DR GO; GO:0005525; F:GTP binding; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR001019; Gprotein\_alpha.  
 DR InterPro; IPR000367; Gprotein\_alpha\_S.  
 DR InterPro; IPR011025; GproteinA\_insert.  
 DR Pfam; PF00503; G-alpha; 1.  
 DR PRINTS; PR00318; GPROTEIN\_A.  
 DR PRINTS; PR00443; GPROTEIN\_A.  
 DR SMART; SM00275; G\_alpha; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 375 AA; 44012 MW; B02C195E9191DD52 CRC64;  
  
 Query Match 51.6%; Score 79; DB 2; Length 375;  
 Best Local Similarity 51.6%; Pred. No. 0.022;  
 Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;  
  
 Qy 1 KEARKVSRGIDRMLRDKQKRDLOQTHRLLLLG 31  
 Db 13 REARKVKNQIEQLAKDKQVMRATHRLLLLG 43  
  
 RESULT 9  
 Q4SQ90 TETNG  
 ID Q4SQ90 TETNG PRELIMINARY; PRT; 485 AA.  
 AC Q4SQ90;  
 DT 13-SEP-2005 (TReMBLrel. 31, Created)  
 DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)  
 DE Chromosome 4 SCAFI4533, whole genome shotgun sequence.  
 DE (Fragment).  
 GN ORFNames=GSTENG00014468001;  
 OS Tetraodon nigroviridis (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 OX NCBI\_TaxID=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
 RA Mauceli E., Bouteau L., Fischer C., Ozouf-Coataz C., Bernot A.,  
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
 RA Dasilva C., Sjuloubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Sublin C., Castelli V., Katinka M., Vacherie B.,  
 RA Biemont C., Skalli Z., Cattolico L., Poulin J., De Berardinis V.,  
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,  
 RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,  
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
 RA Lindblad-Toh K., Birren B., Nuebaum C., Kahn D., Robinson-Rechavi M.,  
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,  
 RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;  
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 RT the early vertebrate proto-karyotype."  
 RL Nature 431:946-957(2004).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Genoscope; Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; CAAE01014533; CAF97192.1; -; Genomic DNA.  
 .FT NON\_TER 1  
 SQ SEQUENCE 485 AA; 55733 MW; B90456AC53393CB2 CRC64;  
  
 Query Match 51.0%; Score 78; DB 2; Length 485;  
 Best Local Similarity 51.6%; Pred. No. 0.04;  
 Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;  
  
 Qy 1 KEARKVSRGIDRMLRDKQKRDLOQTHRLLLLG 31  
 Db 13 REARKVKNQIEQLAKDKQVMRATHRLLLLG 43







```

SQ  SEQUENCE      359 AA;  42142 MW;  6F69C4F617DFA7C7 CRC64;

Query Match          49.7%;  Score 76;  DB 2;  Length 359;
Best Local Similarity 51.6%;  Pred. No. 0.056;
Matches 16;  Conservative 7;  Mismatches 8;  Indels 0;  Gaps 0;

QY  1 KEARKVSRGIDRMRLDQKRDQLQOTHRLLLG 31
    |||||::: ||:: |||||
DB  16 KEARRINDEIERQLRRDKEDARRELKLLLG 46

RESULT 15
QSF3B5_CHICK
ID  QSF3B5_CHICK PRELIMINARY;  PRT;  359 AA.
AC  QSF3B5;
DT  10-MAY-2005 (TrEMBLrel. 30, Created)
DT  10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE  10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE  Hypothetical protein.
GN  ORFNames=RCJMB04_23b22;
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=CB; TISSUE=Bursa;
RA  Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA  Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA  Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT  "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT  genefunction analysis."
RL  Genome Biol. 6:R6-R6(2005).
DR  EMBL; AJ851735; CAH6S369.1; -; mRNA.
DR  InterPro; IPR011025; Gproteina_insert.
DR  InterPro; IPR001019; Gproteina_alpha.
DR  InterPro; IPR000654; Gproteina_alpha_Q.
DR  Pfam; PF00503; G-alpha; 1.
DR  PRINTS; PR00318; GPROTEINA.
DR  PRINTS; PR00442; GPROTEINAQ.
DR  ProDom; PD000281; Gproteina_alpha; 1.
DR  SMART; SM00275; G_alpha; 1.
KW  Hypothetical protein.
SQ  SEQUENCE      359 AA;  42234 MW;  CFICE9B63E7153D3 CRC64;

Query Match          49.7%;  Score 76;  DB 2;  Length 359;
Best Local Similarity 51.6%;  Pred. No. 0.056;
Matches 16;  Conservative 7;  Mismatches 8;  Indels 0;  Gaps 0;

QY  1 KEARKVSRGIDRMRLDQKRDQLQOTHRLLLG 31
    |||||::: ||:: |||||
DB  16 KEARRINDEIERQLRRDKEDARRELKLLLG 46

Search completed: March 2, 2006, 19:37:30
Job time : 21.3988 secs

```

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:30:47 ; Search time 2.409 Seconds  
(without alignments)  
1063.905 Million cell updates/sec

Title: US-10-618-320A-1\_COPY\_96\_126

Perfect score: 153

Sequence: 1 KEARKVSRGIDRLMDKRDQLQTHRLLLIG 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	49.7	353	2	US-09-984-292-2
2	76	49.7	353	2	US-09-984-292-4
3	76	49.7	353	2	US-09-984-292-5
4	76	49.7	353	2	US-09-984-292-16
5	76	49.7	359	2	US-09-984-292-1
6	76	49.7	359	2	US-09-984-292-3
7	76	49.7	359	2	US-09-984-292-10
8	76	49.7	359	2	US-09-984-292-15
9	76	49.7	359	2	US-09-984-292-20
10	76	49.7	359	2	US-09-984-292-37
11	71	46.4	353	2	US-09-984-292-6
12	71	46.4	353	2	US-09-984-292-8
13	71	46.4	353	2	US-09-984-292-14
14	71	46.4	353	2	US-09-984-292-17
15	71	46.4	353	2	US-09-984-292-18
16	71	46.4	353	2	US-09-984-292-19
17	71	46.4	353	2	US-09-984-292-24
18	71	46.4	357	2	US-09-984-292-7
19	71	46.4	359	2	US-09-984-292-9
20	71	46.4	359	2	US-09-984-292-11
21	71	46.4	359	2	US-09-984-292-12
22	71	46.4	359	2	US-09-984-292-13
23	71	46.4	359	2	US-09-984-292-21
24	71	46.4	359	2	US-09-984-292-22
25	71	46.4	359	2	US-09-984-292-23
26	71	46.4	359	2	US-09-984-292-25
27	71	46.4	359	2	US-09-984-292-26

```

28 70 45.8 781 2 US-09-949-016-9170 Sequence 9170, Ap
29 68 44.4 351 2 US-09-949-016-7853 Sequence 7853, Ap
30 67 43.8 388 2 US-09-949-016-9025 Sequence 9025, Ap
31 65 42.5 355 2 US-09-492-028-2 Sequence 2, Appli
32 64 41.8 65 1 US-08-464-531-83 Sequence 83, Appli
33 64 41.8 65 1 US-08-461-598-83 Sequence 83, Appli
34 64 41.8 65 2 US-08-322-137-83 Sequence 83, Appli
35 64 41.8 380 2 US-08-307-896-1 Sequence 1, Appli
36 64 41.8 394 2 US-09-442-349A-106 Sequence 106, App
37 64 41.8 394 4 PCT-US95-11808-1 Sequence 1, Appli
38 64 41.8 775 2 US-09-513-838-6 Sequence 6, Appli
39 64 41.8 869 2 US-10-314-048A-100 Sequence 100, App
40 64 41.8 926 2 US-10-314-048A-104 Sequence 104, App
41 64 41.8 1181 2 US-09-826-509-587 Sequence 587, App
42 63 41.2 381 2 US-09-949-016-11694 Sequence 11694, A
43 62 40.5 355 2 US-09-442-349A-107 Sequence 107, App
44 62 40.5 358 2 US-09-949-016-11528 Sequence 11528, A
45 62 40.5 365 2 US-09-442-349A-108 Sequence 108, App

```

#### ALIGNMENTS

##### RESULT 1

US-09-984-292-2

; Sequence 2, Application US/09984292

; Patent No. 6818747

; GENERAL INFORMATION:

; APPLICANT: YAO, YONG

; APPLICANT: XU, HONG

; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF CHEMOSENSORY RECEPTORS

; FILE REFERENCE: 078003-0280649

; CURRENT APPLICATION NUMBER: US/09/984,292

; CURRENT FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: 60/243,770

; PRIOR FILING DATE: 2000-10-30

; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 353

; TYPE: PRT

; ORGANISM: Mus sp.

US-09-984-292-2

Query Match 49.7%; Score 76; DB 2; Length 353;

Best Local Similarity 51.8%; Pred. No. 0.0012; Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRLMDKRDQLQTHRLLLIG 31

DB 10 KEARRINDEIERQLRRDKRDARRELKLLLLG 40

##### RESULT 2

US-09-984-292-4

; Sequence 4, Application US/09984292

; Patent No. 6818747

; GENERAL INFORMATION:

; APPLICANT: YAO, YONG

; APPLICANT: XU, HONG

; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF CHEMOSENSORY RECEPTORS

; FILE REFERENCE: 078003-0280649

; CURRENT APPLICATION NUMBER: US/09/984,292

; CURRENT FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: 60/243,770

; PRIOR FILING DATE: 2000-10-30

; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4



```

; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-984-292-4

```

Query Match 49.7%; Score 76; DB 2; Length 353;  
Best Local Similarity 51.6%; Pred. No. 0.0012;  
Matches 16; Conservative 7; Mismatches 8; Indels

Qy 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31  
|||||:::|:| ||| ::|:|:|

Db 10 KEARRINDEIERQLRRDRKRDARRELKLLLG 40

```

RESULT 3
US-09-984-292-5
; Sequence 5, Application US/09984292
;
; Patent No. 6818747
; GENERAL INFORMATION:
;
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
;
; TITLE OF INVENTION: G-ALPHA-Q PROT
; TITLE OF INVENTION: ANALYSIS AND
; TITLE OF INVENTION: CHEMOSENSORY
; FILE REFERENCE: 078003-0280649
;
; CURRENT APPLICATION NUMBER: US/09/
; CURRENT FILING DATE: 2001-10-29
;
; PRIOR APPLICATION NUMBER: 60/243,7
; PRIOR FILING DATE: 2000-10-30
;
; NUMBER OF SEQ ID NOS: 42
;
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
;
; LENGTH: 353
;
; TYPE: PRT
;
; ORGANISM: Mus sp.
US-09-984-292-5

```

Query Match 49.7%; Score 76; DB 2; Length 353;  
Best Local Similarity 51.6%; Pred. No. 0.0012;  
Matches 16; Conservative 7; Mismatches 8; Indels

Qy 1 KEARKVSRGIDRWMLRDQKRDQLQOTHRLLLLG 31  
 ||||::: |:| || ||| :: :|||  
 Db 10 KEARRINDEIERQLRRDRDARRELKLLLG 40

```

RESULT 4
US-09-984-292-16
; Sequence 16, Application US/09984292
; Patent No. 6818747
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: US/09/984,292
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-292-16

```

Query Match 49.7%; Score 76; DB 2; Length 353;  
Best Local Similarity 51.6%; Pred. No. 0.0012;  
Matches 16; Conservative 7; Mismatches 8; Indels

**Qy** 1 KEARKVSRGIDRWLRDQKRDQLQTHRLLLLG 31  
|||||::: |: | | ||| :: :|||  
**Db** 10 KEARRINDEIERQLRRDKRDARRELKLLLG 40

RESULT 5  
 US-09-984-292-1  
 ; Sequence 1, Application US/09984292  
 ; Patent No. 6818747  
 ; GENERAL INFORMATION:  
 ; APPLICANT: YAO, YONG  
 ; APPLICANT: XU, HONG  
 ; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
 ; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
 ; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS  
 ; FILE REFERENCE: 078003-0280649  
 ; CURRENT APPLICATION NUMBER: US/09/984,292  
 ; CURRENT FILING DATE: 2001-10-29  
 ; PRIOR APPLICATION NUMBER: 60/243,770  
 ; PRIOR FILING DATE: 2000-10-30  
 ; NUMBER OF SEQ ID NOS: 42  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 359  
 ; TYPE: PRT  
 ; ORGANISM: Mus sp.  
 US-09-984-292-1

Query Match 49.7%; Score 76; DB 2; Length 359;  
Best Local Similarity 51.6%; Pred. No. 0.0012;  
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KEARKVSRGIDRMRLDQKRDLOQTHRLLLLG 31  
|||:::|:|:| |||:::|:|:|  
Db 16 KEARRINDEIERQLRRDKRDARRELKLLLG 46

```

RESULT 6
US-09-984-292-3
; Sequence 3, Application US/09984292
;
; PATENT INFORMATION:
; GENERAL INFO: 6818747
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
;
; TITLE OF INVENTION: G-ALPHA-Q PROT
; TITLE OF INVENTION: ANALYSIS AND
; TITLE OF INVENTION: CHEMOSENSORY I
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: US/09/
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,7
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-984-292-3

```

Query Match 49.7%; Score 76; DB 2; Length 359;  
Best Local Similarity 51.6%; Pred. No. 0.0012;  
Matches 16: Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRWLRDQKRD LQTHRLLLG 31  
 ||||::: |:| | | ||| : ||||  
 Db 16 KEARRINDEIERQLRRDKRDARELKL LLLG 46

RESULT 7  
US-09-984-292-10  
; Sequence 10, Application US/09984292  
; Patent No. 6818747

```

; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-292-20

Query Match          49.7%; Score 76; DB 2; Length 359;
Best Local Similarity 51.6%; Pred.No. 0.0012;
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY      1 KEARKVSRGIDRMRLDQKEDLQOQTHRLLLLG 31
      ||||::: ||::: ||||::: ||||:::
DB      16 KEARRINDEIERQLRRDKRDARRELKLLLG 46

RESULT 10
US-09-984-292-37
; Sequence 37, Application US/09984292
; Patent No. 6818747
; GENERAL INFORMATION:
; APPLICANT: YAO, HONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: US/09/984,292
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-292-37

Query Match          49.7%; Score 76; DB 2; Length 359;
Best Local Similarity 51.6%; Pred.No. 0.0012;
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY      1 KEARKVSRGIDRMRLDQKEDLQOQTHRLLLLG 31
      ||||::: ||::: ||||::: ||||:::
DB      16 KEARRINDEIERQLRRDKRDARRELKLLLG 46

RESULT 11
US-09-984-292-6
; Sequence 6, Application US/09984292
; Patent No. 6818747
; GENERAL INFORMATION:
; APPLICANT: YAO, HONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: US/09/984,292
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-984-292-6

```

Search completed: March 2, 2006, 19:31:50  
Job time : 3.409 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using 'sw model

Run on: March 2, 2006, 19:32:31 ; Search time 14.137 Seconds  
(without alignments)  
916.227 Million cell updates/sec

Title: US-10-618-320A-1\_COPY\_96\_126

Perfect score: 153

Sequence: 1 KEARKVSRGIDRLRDKEDLQOTHRLLLG 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*

3: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*

4: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

5: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep.\*

6: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153	100.0	458	5	US-10-732-923-8011
2	153	100.0	461	5	US-10-450-763-31595
3	149	97.4	230	5	US-10-732-923-7629
4	95	62.1	351	5	US-10-732-923-8014
5	79	51.6	375	5	US-10-732-923-7864
6	79	51.6	378	5	US-10-732-923-7865
7	79	51.6	385	4	US-10-369-493-4978
8	79	51.6	385	5	US-10-732-923-7867
9	76	49.7	353	3	US-09-984-292-2
10	76	49.7	353	3	US-09-984-292-4
11	76	49.7	353	3	US-09-984-292-5
12	76	49.7	353	3	US-09-984-292-16
13	76	49.7	353	3	US-09-989-497-2
14	76	49.7	353	3	US-09-989-497-4
15	76	49.7	353	3	US-09-989-497-5
16	76	49.7	353	3	US-09-989-497-16
17	76	49.7	353	3	US-09-952-680A-27
18	76	49.7	353	4	US-10-352-843-13
19	76	49.7	353	5	US-10-215-982-27
20	76	49.7	353	5	US-10-732-923-7646
21	76	49.7	353	5	US-10-732-923-7999
22	76	49.7	353	5	US-10-732-923-8036
23	76	49.7	353	5	US-10-931-080B-5
24	76	49.7	353	5	US-10-952-773-2
25	76	49.7	353	5	US-10-952-773-4
26	76	49.7	353	5	US-10-952-773-16
27	76	49.7	359	3	US-09-984-292-1

28	76	49.7	359	3	US-09-984-292-3	Sequence 3, Appli
29	76	49.7	359	3	US-09-984-292-10	Sequence 10, Appl
30	76	49.7	359	3	US-09-984-292-15	Sequence 15, Appl
31	76	49.7	359	3	US-09-984-292-20	Sequence 20, Appl
32	76	49.7	359	3	US-09-984-292-37	Sequence 37, Appl
33	76	49.7	359	3	US-09-989-497-1	Sequence 1, Appli
34	76	49.7	359	3	US-09-989-497-3	Sequence 3, Appli
35	76	49.7	359	3	US-09-989-497-10	Sequence 10, Appl
36	76	49.7	359	3	US-09-989-497-15	Sequence 15, Appl
37	76	49.7	359	3	US-09-989-497-20	Sequence 20, Appl
38	76	49.7	359	3	US-09-989-497-37	Sequence 37, Appl
39	76	49.7	359	4	US-10-059-266B-2	Sequence 2, Appli
40	76	49.7	359	5	US-10-732-923-7647	Sequence 7647, Ap
41	76	49.7	359	5	US-10-732-923-7961	Sequence 7961, Ap
42	76	49.7	359	5	US-10-732-923-7995	Sequence 7995, Ap
43	76	49.7	359	5	US-10-732-923-8000	Sequence 8000, Ap
44	76	49.7	359	5	US-10-732-923-8002	Sequence 8002, Ap
45	76	49.7	359	5	US-10-732-923-8037	Sequence 8037, Ap

## ALIGNMENTS

RESULT 1  
US-10-732-923-8011  
; Sequence 8011, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 8011  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-732-923-8011

Query Match 100.0%; Score 153; DB 5; Length 458;  
Best Local Similarity 100.0%; Pred. No. 6.4e-13;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KEARKVSRGIDRLRDKEDLQOTHRLLLG 31  
Db 96 KEARKVSRGIDRLRDKEDLQOTHRLLLG 126

RESULT 2  
US-10-450-763-31595  
; Sequence 31595, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 31595  
; LENGTH: 461  
; TYPE: PRT  
; ORGANISM: Homo sapiens



Query Match 51.6%; Score 79; DB 4; Length 385;  
Best Local Similarity 51.6%; Pred. No. 0.012;  
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31  
Db 16 REARKVNKQIEEQLAKDKQVWRATHRLLLG 46

## RESULT 8

US-10-732-923-7867  
; Sequence 7867, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 7867  
; LENGTH: 385  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-10-732-923-7867

Query Match 51.6%; Score 79; DB 5; Length 385;  
Best Local Similarity 51.6%; Pred. No. 0.012;  
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31  
Db 16 REARKVNKQIEEQLAKDKQVWRATHRLLLG 46

## RESULT 9

US-09-984-292-2  
; Sequence 2, Application US/09984292  
; Patent No. US20020128433A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280649  
; CURRENT APPLICATION NUMBER: US/09/984,292  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,770  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-984-292-2

Query Match 49.7%; Score 76; DB 3; Length 353;  
Best Local Similarity 51.6%; Pred. No. 0.029;  
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31  
Db 10 KEARRINDEIERQLRRDKRDARRELKLLLG 40

## RESULT 10

US-09-984-292-4  
; Sequence 4, Application US/09984292

; Patent No. US20020128433A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280649  
; CURRENT APPLICATION NUMBER: US/09/984,292  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,770  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-984-292-4

Query Match 49.7%; Score 76; DB 3; Length 353;  
Best Local Similarity 51.6%; Pred. No. 0.029;  
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31  
Db 10 KEARRINDEIERQLRRDKRDARRELKLLLG 40

## RESULT 11

US-09-984-292-5  
; Sequence 5, Application US/09984292  
; Patent No. US20020128433A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280649  
; CURRENT APPLICATION NUMBER: US/09/984,292  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,770  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-984-292-5

Query Match 49.7%; Score 76; DB 3; Length 353;  
Best Local Similarity 51.6%; Pred. No. 0.029;  
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31  
Db 10 KEARRINDEIERQLRRDKRDARRELKLLLG 40

## RESULT 12

US-09-984-292-16  
; Sequence 16, Application US/09984292  
; Patent No. US20020128433A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280649  
; CURRENT APPLICATION NUMBER: US/09/984,292

```
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-292-16

Query Match          49.7%; Score 76; DB 3; Length 353;
Best Local Similarity 51.6%; Pred. No. 0.029;
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31
    |||||::: |:| || ||| :: : |||||
Db 10 KEARRINDEIERQLRRDKRDARRELKLLLG 40

RESULT 13
US-09-989-497-2
; Sequence 2, Application US/09989497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280735
; CURRENT APPLICATION NUMBER: US/09/989,497
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/984,292
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-989-497-2

Query Match          49.7%; Score 76; DB 3; Length 353;
Best Local Similarity 51.6%; Pred. No. 0.029;
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31
    |||||::: |:| || ||| :: : |||||
Db 10 KEARRINDEIERQLRRDKRDARRELKLLLG 40

RESULT 14
US-09-989-497-4
; Sequence 4, Application US/09989497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280735
; CURRENT APPLICATION NUMBER: US/09/989,497
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/984,292
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-989-497-4

Query Match          49.7%; Score 76; DB 3; Length 353;
Best Local Similarity 51.6%; Pred. No. 0.029;
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31
    |||||::: |:| || ||| :: : |||||
Db 10 KEARRINDEIERQLRRDKRDARRELKLLLG 40

Search completed: March 2, 2006, 19:41:16
Job time : 14.137 secs
```

```
; SEQ ID NO 4
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-989-497-4

Query Match          49.7%; Score 76; DB 3; Length 353;
Best Local Similarity 51.6%; Pred. No. 0.029;
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31
    |||||::: |:| || ||| :: : |||||
Db 10 KEARRINDEIERQLRRDKRDARRELKLLLG 40

RESULT 15
US-09-989-497-5
; Sequence 5, Application US/09989497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280735
; CURRENT APPLICATION NUMBER: US/09/989,497
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/984,292
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-989-497-5

Query Match          49.7%; Score 76; DB 3; Length 353;
Best Local Similarity 51.6%; Pred. No. 0.029;
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLDQKRDQLQOQTHRLLLG 31
    |||||::: |:| || ||| :: : |||||
Db 10 KEARRINDEIERQLRRDKRDARRELKLLLG 40

Search completed: March 2, 2006, 19:41:16
Job time : 14.137 secs
```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 19:37:27 ; Search time 1.07771 Seconds  
(without alignments)  
575.298 Million cell updates/sec

Title: US-10-618-320A-1\_COPY\_96\_126

Perfect score: 153

Sequence: 1 KEARKVSRGIDRMRLDKEDLQOTHRLLLLG 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 135339 seqs, 20000136 residues

Total number of hits satisfying chosen parameters: 135339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*

7: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153	100.0	458	6	US-10-618-320A-1
2	149	97.4	448	6	US-10-618-320A-25
3	149	97.4	450	6	US-10-618-320A-26
4	76	49.7	859	7	US-11-053-100-52
5	76	49.7	1309	7	US-11-053-100-53
6	71	46.4	353	7	US-11-060-023-2
7	71	46.4	353	7	US-11-060-023-4
8	71	46.4	353	7	US-11-060-023-6
9	71	46.4	353	7	US-11-060-023-8
10	71	46.4	353	7	US-11-060-023-12
11	71	46.4	353	7	US-11-060-023-14
12	71	46.4	359	7	US-11-060-023-16
13	71	46.4	359	7	US-11-060-023-11
14	71	46.4	359	7	US-11-060-023-13
15	71	46.4	360	7	US-11-060-023-17
16	68	44.4	353	7	US-11-060-023-15
17	64	41.8	394	6	US-10-821-234-1626
18	64	41.8	864	7	US-11-053-100-58
19	58	37.9	374	7	US-11-228-364-4
20	54	35.3	374	7	US-11-060-023-10
21	54	35.3	374	7	US-11-228-364-2
22	54	35.3	374	7	US-11-169-041-188
23	53	34.6	641	6	US-10-821-234-1519
24	53	34.6	641	6	US-10-491-096-189
25	51	33.3	646	6	US-10-491-096-190

26	49	32.0	337	7	US-11-087-099-6276	Sequence 6276, Ap
27	46	30.1	452	7	US-11-024-959-382	Sequence 382, App
28	45	29.4	408	6	US-10-821-234-1100	Sequence 1100, Ap
29	45	29.4	411	7	US-11-072-512-3452	Sequence 3452, Ap
30	45	29.4	763	6	US-10-821-234-1619	Sequence 1619, Ap
31	45	29.4	915	6	US-10-821-234-1514	Sequence 1514, Ap
32	45	29.4	915	6	US-10-995-561-1003	Sequence 1003, Ap
33	45	29.4	917	6	US-10-995-561-1000	Sequence 1000, Ap
34	45	29.4	940	6	US-10-995-561-1004	Sequence 1004, Ap
35	45	29.4	969	6	US-10-995-561-1001	Sequence 1001, Ap
36	45	29.4	971	6	US-10-995-561-998	Sequence 998, App
37	45	29.4	994	6	US-10-995-561-997	Sequence 997, App
38	44.5	29.1	299	6	US-10-362-772-2	Sequence 2, Appli
39	44	28.8	250	6	US-10-453-372-938	Sequence 938, App
40	44	28.8	452	7	US-11-024-959-451	Sequence 451, App
41	44	28.8	625	7	US-11-072-512-3210	Sequence 3210, Ap
42	44	28.8	2087	7	US-11-075-185-28	Sequence 28, Appl
43	43.5	28.4	250	7	US-11-087-099-8190	Sequence 8190, Ap
44	43	28.1	246	7	US-11-232-406A-14	Sequence 14, Appl
45	43	28.1	250	6	US-10-453-372-942	Sequence 942, App

## ALIGNMENTS

### RESULT 1

US-10-618-320A-1

; Sequence 1, Application US/10618320A

; Publication No. US20050260595A1

; GENERAL INFORMATION:

; APPLICANT: Sumitomo Chemical Company Limited

; TITLE OF INVENTION: NOVEL G PROTEINS, POLYNUCLEOTIDE ENCODING THE SAME AND UTILIZATION

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/618,320A

; PRIOR FILING DATE: 2003-07-11

; PRIOR APPLICATION NUMBER: JP 2002/206841

; PRIOR FILING DATE: 2002-07-16

; PRIOR APPLICATION NUMBER: JP 2002/367778

; PRIOR FILING DATE: 2002-12-19

; PRIOR APPLICATION NUMBER: JP 2003/095955

; PRIOR FILING DATE: 2003-03-31

; NUMBER OF SEQ ID NOS: 34

; SEQ ID NO 1

; LENGTH: 458

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-618-320A-1

Query Match 100.0%; Score 153; DB 6; Length 458;  
Best Local Similarity 100.0%; Pred. No. 6e-15;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	KEARKVSRGIDRMRLDKEDLQOTHRLLLLG 31
Db	96	KEARKVSRGIDRMRLDKEDLQOTHRLLLLG 126

### RESULT 2

US-10-618-320A-25

; Sequence 25, Application US/10618320A

; Publication No. US20050260595A1

; GENERAL INFORMATION:

; APPLICANT: Sumitomo Chemical Company Limited

; TITLE OF INVENTION: NOVEL G PROTEINS, POLYNUCLEOTIDE ENCODING THE SAME AND UTILIZATION

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/618,320A

; PRIOR FILING DATE: 2003-07-11

; PRIOR APPLICATION NUMBER: JP 2002/206841

; PRIOR FILING DATE: 2002-07-16

; PRIOR APPLICATION NUMBER: JP 2002/367778

; PRIOR FILING DATE: 2002-12-19

; PRIOR APPLICATION NUMBER: JP 2003/095955

; PRIOR FILING DATE: 2003-03-31

```
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 25
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-618-320A-25

Query Match      97.4%; Score 149; DB 6; Length 448;
Best Local Similarity 96.8%; Pred. No. 2.3e-14;
Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLRDQKRDLOQTHRLLLG 31
Db 86 KEARKVSRGIDRMRLREQKRDLOQTHRLLLG 116

RESULT 3
US-10-618-320A-26
; Sequence 26, Application US/10618320A
; Publication No. US20050260595A1
; GENERAL INFORMATION:
; APPLICANT: Sumitomo Chemical Company Limited
; TITLE OF INVENTION: NOVEL G PROTEINS, POLYNUCLEOTIDE ENCODING THE SAME AND UTILIZATION
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/618,320A
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: JP 2002/206841
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: JP 2002/367778
; PRIOR FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: JP 2003/095955
; PRIOR FILING DATE: 2003-03-31
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 26
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-618-320A-26

Query Match      97.4%; Score 149; DB 6; Length 450;
Best Local Similarity 96.8%; Pred. No. 2.3e-14;
Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLRDQKRDLOQTHRLLLG 31
Db 88 KEARKVSRGIDRMRLREQKRDLOQTHRLLLG 118

RESULT 4
US-11-053-100-52
; Sequence 52, Application US/11053100
; Publication No. US2005025554A1
; GENERAL INFORMATION:
; APPLICANT: CHILKOTI, Ashutosh
; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION
; FILE REFERENCE: 4176-101 CIP
; CURRENT APPLICATION NUMBER: US/11/053,100
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US 09/812,382
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,659
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 52
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(859)

US-11-060-023-2
; Sequence 2, Application US/11060023
; Publication No. US20050255531A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pharma Deutschland GmbH
; TITLE OF INVENTION: Process for identifying modulators of G protein coupled
; FILE REFERENCE: AVE D-2000/A033 englisch
; CURRENT APPLICATION NUMBER: US/11/060,023
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: US/09/899,295
; PRIOR FILING DATE: 2003-07-06
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-060-023-2

Query Match      46.4%; Score 71; DB 7; Length 353;
Best Local Similarity 48.4%; Pred. No. 0.0047;
Matches 15; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
```

```
; OTHER INFORMATION: pET15b-SD1-ELP1-90-throm-G protein alpha Q
US-11-053-100-52

Query Match      49.7%; Score 76; DB 7; Length 859;
Best Local Similarity 51.6%; Pred. No. 0.0026;
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLRDQKRDLOQTHRLLLG 31
Db 516 KEARRINDEIERQLARDKRDARRELKLLLLG 546

RESULT 5
US-11-053-100-53
; Sequence 53, Application US/11053100
; Publication No. US2005025554A1
; GENERAL INFORMATION:
; APPLICANT: CHILKOTI, Ashutosh
; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION
; FILE REFERENCE: 4176-101 CIP
; CURRENT APPLICATION NUMBER: US/11/053,100
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US 09/812,382
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,659
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 53
; LENGTH: 1309
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1309)
; OTHER INFORMATION: pET15b-SD1-ELP1-180-throm-G protein alpha Q
US-11-053-100-53

Query Match      49.7%; Score 76; DB 7; Length 1309;
Best Local Similarity 51.6%; Pred. No. 0.0043;
Matches 16; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KEARKVSRGIDRMRLRDQKRDLOQTHRLLLG 31
Db 966 KEARRINDEIERQLARDKRDARRELKLLLLG 996

RESULT 6
US-11-060-023-2
; Sequence 2, Application US/11060023
; Publication No. US20050255531A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pharma Deutschland GmbH
; TITLE OF INVENTION: Process for identifying modulators of G protein coupled
; FILE REFERENCE: AVE D-2000/A033 englisch
; CURRENT APPLICATION NUMBER: US/11/060,023
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: US/09/899,295
; PRIOR FILING DATE: 2003-07-06
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-060-023-2

Query Match      46.4%; Score 71; DB 7; Length 353;
Best Local Similarity 48.4%; Pred. No. 0.0047;
Matches 15; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
```

; GERMANY INFORMATION.  
 ; APPLICANT: Aventis Pharma Deutschland GmbH  
 ; TITLE OF INVENTION: Process for identifying modulators of G protein coupled

Query Match	46.4%;	Score 71;	DB 7;	Length 353;	
Best Local Similarity	48.4%;	Pred. No. 0.0047;			
Matches	15;	Conservative	8;	Mismatches	8; Indels 0; Gaps 0;
QY	1	KEARKVSRGIDRMLRDQKRDLDQOQTHRLLLG	31		
DB	10	KEARRINDEIERHVRDKRDARRELKLLLG	40		
RESULT 12					
US-11-060-023-16					
Sequence 16,	Application US/11060023				
Publication No.	US20050255531A1				
GENERAL INFORMATION:					
APPLICANT:	Aventis Pharma Deutschland GmbH				
TITLE OF INVENTION:	Process for identifying modulators of G protein coupled				
TITLE OF INVENTION:	receptors				
FILE REFERENCE:	AVE D-2000/A033 englisch				
CURRENT APPLICATION NUMBER:	US/11/060,023				
CURRENT FILING DATE:	2005-02-17				
PRIOR APPLICATION NUMBER:	US/09/899,295				
PRIOR FILING DATE:	2003-07-06				
NUMBER OF SEQ ID NOS:	17				
SOFTWARE:	PatentIn Ver. 2.1				
SEQ ID NO 16					
LENGTH:	353				
TYPE:	PRT				
ORGANISM:	Mus musculus				
US-11-060-023-16					
Query Match	46.4%;	Score 71;	DB 7;	Length 353;	
Best Local Similarity	48.4%;	Pred. No. 0.0047;			
Matches	15;	Conservative	8;	Mismatches	8; Indels 0; Gaps 0;
QY	1	KEARKVSRGIDRMLRDQKRDLDQOQTHRLLLG	31		
DB	10	KEARRINDEIERHVRDKRDARRELKLLLG	40		
RESULT 13					
US-11-060-023-11					
Sequence 11,	Application US/11060023				
Publication No.	US20050255531A1				
GENERAL INFORMATION:					
APPLICANT:	Aventis Pharma Deutschland GmbH				
TITLE OF INVENTION:	Process for identifying modulators of G protein coupled				
TITLE OF INVENTION:	receptors				
FILE REFERENCE:	AVE D-2000/A033 englisch				
CURRENT APPLICATION NUMBER:	US/11/060,023				
CURRENT FILING DATE:	2005-02-17				
PRIOR APPLICATION NUMBER:	US/09/899,295				
PRIOR FILING DATE:	2003-07-06				
NUMBER OF SEQ ID NOS:	17				
SOFTWARE:	PatentIn Ver. 2.1				
SEQ ID NO 11					
LENGTH:	359				
TYPE:	PRT				
ORGANISM:	Mus musculus				
US-11-060-023-11					
Query Match	46.4%;	Score 71;	DB 7;	Length 359;	
Best Local Similarity	48.4%;	Pred. No. 0.0048;			
Matches	15;	Conservative	8;	Mismatches	8; Indels 0; Gaps 0;
QY	1	KEARKVSRGIDRMLRDQKRDLDQOQTHRLLLG	31		
DB	16	KEARRINDEIERHVRDKRDARRELKLLLG	46		
RESULT 14					
US-11-060-023-13					
Sequence 13,	Application US/11060023				
Publication No.	US20050255531A1				
GENERAL INFORMATION:					
APPLICANT:	Aventis Pharma Deutschland GmbH				
TITLE OF INVENTION:	Process for identifying modulators of G protein coupled				
TITLE OF INVENTION:	receptors				
FILE REFERENCE:	AVE D-2000/A033 englisch				
CURRENT APPLICATION NUMBER:	US/11/060,023				
CURRENT FILING DATE:	2005-02-17				
PRIOR APPLICATION NUMBER:	US/09/899,295				
PRIOR FILING DATE:	2003-07-06				
NUMBER OF SEQ ID NOS:	17				
SOFTWARE:	PatentIn Ver. 2.1				
SEQ ID NO 11					
LENGTH:	359				
TYPE:	PRT				
ORGANISM:	Mus musculus				
US-11-060-023-13					
Query Match	46.4%;	Score 71;	DB 7;	Length 359;	
Best Local Similarity	48.4%;	Pred. No. 0.0048;			
Matches	15;	Conservative	8;	Mismatches	8; Indels 0; Gaps 0;
QY	1	KEARKVSRGIDRMLRDQKRDLDQOQTHRLLLG	31		
DB	16	KEARRINDEIERHVRDKRDARRELKLLLG	46		
RESULT 15					
US-11-060-023-17					
Sequence 17,	Application US/11060023				
Publication No.	US20050255531A1				
GENERAL INFORMATION:					
APPLICANT:	Aventis Pharma Deutschland GmbH				
TITLE OF INVENTION:	Process for identifying modulators of G protein coupled				
TITLE OF INVENTION:	receptors				
FILE REFERENCE:	AVE D-2000/A033 englisch				
CURRENT APPLICATION NUMBER:	US/11/060,023				
CURRENT FILING DATE:	2005-02-17				
PRIOR APPLICATION NUMBER:	US/09/899,295				
PRIOR FILING DATE:	2003-07-06				
NUMBER OF SEQ ID NOS:	17				
SOFTWARE:	PatentIn Ver. 2.1				
SEQ ID NO 17</					